

STIC-Biotech/ChemLib

90905

From:
Sent:
To:
Subject:

Chan, Christina
Monday, April 07, 2003 4:17 PM
Nguyen, Quang (AU1632); STIC-Biotech/ChemLib
RE: RUSH sequence search request for 09/761640

Please rush. Thanks Chris

Chris Chan
TC 1600 New Hire Training Coordinator and SPE 1644
308-3973
CM-1, 9B19

CRFE

-----Original Message-----

From: **Nguyen, Quang (AU1632)**
Sent: Monday, April 07, 2003 4:14 PM
To: Chan, Christina
Subject: RUSH sequence search request for 09/761640

I would like to request for a RUSH sequence search for the above application, because I need a count for this bi-week.

Please search:

A nucleic acid encoding the amino acid of SEQ ID NO:4;
A nucleic acid of SEQ ID NO:1; and
A nucleic acid of SEQ ID NO:3;
against commercial, issued and pending US patent application databases.

I am in AU1636, my mailbox is in CM1-11E12.

THANK YOU.

Edward Hart
Technical Info. Specialist
STIC/Biotech
April 7, 2003 Tel: 305-9203

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 4/8/03
Date Completed: 4/14/03
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:

NA Sequences: 9
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

reverted
NA

VENDOR/COST (where applic.)

STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: 1#
WWW/Internet: _____
Other (specify): _____

XX	(PEKE) PE CORP NY.
PA	
XX	
PI	Wei M, Ketchum KA, Di Francesco V, Beasley EM;
XX	
DR	WPI; 2002-575237/61.
DR	P-PSDB; ABP51653.
XX	
PT	Novel isolated human phosphatase peptide useful for treating disorder
PT	characterized by absence of, inappropriate or unwanted expression of
PT	the phosphatase protein, and as immunogens to raise antibodies -
PS	
XX	
CC	Claim 1; Fig 1A; 85pp; English.
XX	
CC	The present invention describes an isolated human phosphatase peptide
CC	(I) which can be used for identifying a modulator of (I) by contacting
CC	(I) with an agent and determining if the agent has modulated the
CC	function or activity of (I). (I) is useful for identifying an agent that
CC	binds to (I), by contacting (I) with an agent and assaying the contacted
CC	mixture to determine whether a complex is formed with the agent bound
CC	(I). The human phosphatases from the present invention are mitogen
CC	activated protein (MAP) kinase phosphatases. These human MAP kinase
CC	phosphatases are located on chromosome 11. (I) and the polynucleotide
CC	sequences encoding (I) can be used in gene therapy. The present sequence
CC	encodes human MAP kinase phosphatase splice form 1 from the present
CC	invention.
XX	
SQ	Sequence 2704 BP; 569 A; 874 C; 794 G; 467 T; 0 other;
	Query Match 100.0%; Score 2704; DB 24; Length 2704;
	Best Local Similarity 100.0%; Pred. No. 0;
	Matches 2704; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	1 CGTCTTCCTGTCGTCCGGGATCAGACTGTCGCCGGGTTGAGGAAGCGCCGTGCC 60
DB	1 CGTCTTCCTGTCGTCCGGGATCAGACTGTCGCCGGGTTGAGGAAGCGCCGTGCC 60
OY	61 CGGTGCACCCAGAGTGTCGCGGCGCTGCATCGCCCTGGTCAAGTAGACCCTTGC 120
DB	61 CGGTGCACCCAGAGTGTCGCGGCGCTGCATCGCCCTGGTCAAGTAGACCCTTGC 120
OY	121 CCCCCGGGAGCGGCGCTCCACAGCCGCTGAGGGCCCTGGGACACAGCGGTCCACGAGAAG 180
DB	121 CCCCCGGGAGCGGCGCTCCACAGCCGCTGAGGGCCCTGGGACACAGCGGTCCACGAGAAG 180
OY	181 AGTCACACTCCAGGAAGGCGACAGACTTTGGGGTCTCCGTGGGGCTCTCTGGGACTGCGAG 240
DB	181 AGTCACACTCCAGGAAGGCGACAGACTTTGGGGTCTCCGTGGGGCTCTCTGGGACTGCGAG 240
OY	241 GATGAGGGGACAATATGATGTGACAGAGGCGCATTTTGAGCACACAGAGAAGGCGCCG 300
DB	241 GATGAGGGGACAATATGATGTGACAGAGGCGCATTTTGAGCACACAGAGAAGGCGCCG 300
OY	301 AGTGAAGAGAGGAGCTCCACGCGGACACAGACAGACTTGGGGCAAGAGATCCAGAGTCCCAG 360
DB	301 AGTGAAGAGAGGAGCTCCACGCGGACACAGACAGACTTGGGGCAAGAGATCCAGAGTCCCAG 360
OY	361 AACGAGAGAGAGCGACAGGCGACACTGTCACTTATGCTGACAGCTGCTGAGGCGCGAGAT 420
DB	361 AACGAGAGAGAGCGACAGGCGACACTGTCACTTATGCTGACAGCTGCTGAGGCGCGAGAT 420
OY	421 GACATCGGCTGGCACCCAGCTGAGAGGGAACCCCGGCTCCCGGCTCCGCTACCTGCTG 480
DB	421 GACATCGGCTGGCACCCAGCTGAGAGGGAACCCCGGCTCCCGGCTCCGCTACCTGCTG 480
OY	481 GTAGTTTCTACAGAGAAGGAGAGGCTTGAGGCGAGATGAGAGCGGTCTCTCTGGGCGTG 540
DB	481 GTAGTTTCTACAGAGAAGGAGAGGCTTGAGGCGAGATGAGAGCGGTCTCTCTGGGCGTG 540
OY	541 GATTTCCTGACAGAGACTCCCCAGCTGCACCTTGAGGCTGTGCTTGGCCCTCTGGAGT 600
DB	541 GATTTCCTGACAGAGACTCCCCAGCTGCACCTTGAGGCTGTGCTTGGCCCTCTGGAGT 600

QY	601	GAACCCAGGCTACTAGATGAGACGGGGGCTTCAAGGTGAGCTGCGTGGGCAAGC	660
Db	601	GACACCCAGGGTACTAGATGAGACGGGGGCTTCAAGGTGAGCTGCGTGGGCAAGC	660
QY	661	CGAATCTTCAAGCCATCTCATATCCAGACCAAGTGGGGCACACTCCAGATATTGACACAA	720
Db	661	CGAATCTTCAAGCCATCTCATATCCAGACCAAGTGGGGCACACTCCAGATATTGACACAA	720
QY	721	GCATGTGAGCAGCTCTAGGACAGCGGCTTTGTACCGGGTGGCAGTGGCTTCACTGGGCC	780
Db	721	GCATGTGAGCAGCTCTAGGACAGCGGCTTTGTACCGGGTGGCAGTGGCTTCACTGGGCC	780
QY	781	AGCCACTTACAGGAGAGACTGAATCCGAAACAGAGCTGCTCAATGATGGACGGCTATG	840
Db	781	AGCCACTTACAGGAGAGACTGAATCCGAAACAGAGCTGCTCAATGATGGACGGCTATG	840
QY	841	GCCGACCTGAGGCTCTGCGGGCCCTCCAGGGCCAGACCTGGCGGGTCTTCAGACACAGAG	900
Db	841	GCCGACCTGAGGCTCTGCGGGCCCTCCAGGGCCAGACCTGGCGGGTCTTCAGACACAGAG	900
QY	901	CAGATGAGACAGGGCATCCGCTGTGAGCTGTGAAAAGTTTGGATGTCAGTGACCTGGAG	960
Db	901	CAGATGAGACAGGGCATCCGCTGTGAGCTGTGAAAAGTTTGGATGTCAGTGACCTGGAG	960
QY	961	AGTGTCACTTCCAAAGAGATCCCGCAGAGCTCTGAGAGCTGGCTTGGGGCTCCCCCTCAG	1020
Db	961	AGTGTCACTTCCAAAGAGATCCCGCAGAGCTCTGAGAGCTGGCTTGGGGCTCCCCCTCAG	1020
QY	1021	CAGTACCGTGTACTTATTCGCAACACAGATGCTGCTGGTGGCACACCGGGACCGAGCC	1080
Db	1021	CAGTACCGTGTACTTATTCGCAACACAGATGCTGCTGGTGGCACACCGGGACCGAGCC	1080
QY	1081	TCCCGCATCTTCCCCCACCCTTACCTGGGGCTCAGAGTGGAAACGACAAACCTTGGAGAG	1140
Db	1081	TCCCGCATCTTCCCCCACCCTTACCTGGGGCTCAGAGTGGAAACGACAAACCTTGGAGAG	1140
QY	1141	CTGCAAGAGAACAGAGGTACCCACATCTTGAACATGAGTGGCGGGAGATTGACAACTTCTAC	1200
Db	1141	CTGCAAGAGAACAGAGGTACCCACATCTTGAACATGAGTGGCGGGAGATTGACAACTTCTAC	1200
QY	1201	CCTGAGCCCTTACCTTACACAAATGTCGCGCTCTGGAGTGAAGAGTCCGGCCAGCTGCTG	1260
Db	1201	CCTGAGCCCTTACCTTACACAAATGTCGCGCTCTGGAGTGAAGAGTCCGGCCAGCTGCTG	1260
QY	1261	CCGCACTGGAAGGAGAGCGACACCGCTTCAATTGAGGCTGCAAGAGACAGAGGCACCCACGTG	1320
Db	1261	CCGCACTGGAAGGAGAGCGACACCGCTTCAATTGAGGCTGCAAGAGACAGAGGCACCCACGTG	1320
QY	1321	CTGGTTCACATCGAAGATGGGCTGACCGGCTCAGGCGGCCACAGTGTCTGGCTATGGCATG	1380
Db	1321	CTGGTTCACATCGAAGATGGGCTGACCGGCTCAGGCGGCCACAGTGTCTGGCTATGGCATG	1380
QY	1381	AAGCAGTACGAATGAGCGCTGGACAGAGCCCTGCGGCACGTGACAGAGCTCCGGCCATTC	1440
Db	1381	AAGCAGTACGAATGAGCGCTGGACAGAGCCCTGCGGCACGTGACAGAGCTCCGGCCATTC	1440
QY	1441	GGCCCCCACAACCTTGGCTTCTTGGCGCAGACTTCAGATCTTACAGAGGATCCTTGAACGCC	1500
Db	1441	GGCCCCCACAACCTTGGCTTCTTGGCGCAGACTTCAGATCTTACAGAGGATCCTTGAACGCC	1500
QY	1501	AGAACCTTGAAGGTGTGGGAGAGAAAGTTTGTAGCATGAAGAGACGACAGGACGCC	1560
Db	1501	AGAACCTTGAAGGTGTGGGAGAGAAAGTTTGTAGCATGAAGAGACGACAGGACGCC	1560
QY	1561	GAAAGAAAGCCTGGGCCACGGCCACCTTAAACCTCCGAGGGGTCTAGAGGTCATCAG	1620
Db	1561	GAAAGAAAGCCTGGGCCACGGCCACCTTAAACCTCCGAGGGGTCTAGAGGTCATCAG	1620
QY	1621	TCTTCTGAGAGCCTCTCTTGGAGCTGAGAGACACTCAGAGACAGATGACATGCCAGAGGT	1680
Db	1621	TCTTCTGAGAGCCTCTCTTGGAGCTGAGAGACACTCAGAGACAGATGACATGCCAGAGGT	1680
QY	1681	CTTCTCTTCCACGAGTCTTCACTGAAGAAGCCTTTCGACGCCCTTCCACAGACTTGCAG	1740

```

Db 1681 CTTCTCTCCACGAGCTTTCACATGAGAGAGCTCTGACGCCCTTCCACAGCTTGCAAG 1740
QY 1741 GACCAAGGAGGAGCCACAGAGGTGGAGAGGGGGCTTCAGCTGCTCCCTGAAGTCCCGCACTC 1800
Db 1741 GACCAAGGAGGAGCCACAGAGGTGGAGAGGGGGCTTCAGCTGCTCCCTGAAGTCCCGCACTC 1800
QY 1801 AGTGGTTACCTCCAGAGGAGCTGCGTGGTGGCCAAACCGGAGCCAGGCTTCCAGAGACA 1860
Db 1801 AGTGGTTACCTCCAGAGGAGCTGCGTGGTGGCCAAACCGGAGCCAGGCTTCCAGAGACA 1860
QY 1861 GGAGCAGGAGGAGGAGGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920
Db 1861 GGAGCAGGAGGAGGAGGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920
QY 1921 GAAGGTGGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1980
Db 1921 GAAGGTGGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1980
QY 1981 CACACATGCCACGCTCCCTGACACTGAAAGAGATCCACAACTCTTGAGAGAAACACCC 2040
Db 1981 CACACATGCCACGCTCCCTGACACTGAAAGAGATCCACAACTCTTGAGAGAAACACCC 2040
QY 2041 TCACGCTGTTGGCGACACATTCCTCTCAGCTCCGCCCAATACCCGTCACCTACAGCCTC 2100
Db 2041 TCACGCTGTTGGCGACACATTCCTCTCAGCTCCGCCCAATACCCGTCACCTACAGCCTC 2100
QY 2101 ACCGTCACCCCTGACACTGAGGCTGACCTCCACCCCTGACACTGAGGCTGACCTCC 2160
Db 2101 ACCGTCACCCCTGACACTGAGGCTGACCTCCACCCCTGACACTGAGGCTGACCTCC 2160
QY 2161 TACAGCCTTAAGTCCACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2220
Db 2161 TACAGCCTTAAGTCCACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2220
QY 2221 GATGCTGAGAGGAGAGTGAAGTACCTTTGGGGGCAACAGACACCTAGTTTCATTCTCAA 2280
Db 2221 GATGCTGAGAGGAGAGTGAAGTACCTTTGGGGGCAACAGACACCTAGTTTCATTCTCAA 2280
QY 2281 CTCTAGCCCTGACACACTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2340
Db 2281 CTCTAGCCCTGACACACTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2340
QY 2341 GGTGACGCTCCACAGCCCGCCCTCCCTGACCTCTGCTCTCTCTCCAGTTCAATTCCT 2400
Db 2341 GGTGACGCTCCACAGCCCGCCCTCCCTGACCTCTGCTCTCTCTCCAGTTCAATTCCT 2400
QY 2401 GGAACCAAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2460
Db 2401 GGAACCAAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2460
QY 2461 GCGCGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2520
Db 2461 GCGCGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2520
QY 2521 TCTTGTGCTGTGCTCCAGACCTCTGTGACACAGCCAGATACAGAGGAGCAGAGGCA 2580
Db 2521 TCTTGTGCTGTGCTCCAGACCTCTGTGACACAGCCAGATACAGAGGAGCAGAGGCA 2580
QY 2581 GAGATAGTCTTCTTTTGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2640
Db 2581 GAGATAGTCTTCTTTTGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2640
QY 2641 GTATCTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2700
Db 2641 GTATCTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2700
QY 2701 AAAA 2704
Db 2701 AAAA 2704

```

```

AB073252
ID AB073252 standard; cdna: 2704 BP.
AC AC
XX
XX
30-SEP-2002 (first entry)
DE Human MAP kinase phosphatase splice form 1 cDNA sequence SEQ ID NO:7.
XX
XX Human: phosphatase; mitogen activated protein kinase phosphatase;
KM MAP kinase; enzyme; chromosome 11; single nucleotide polymorphism;
KW SNP; gene; ss.
XX
XX Homo sapiens.
FH Key Location/Qualifiers
FT 5'UTR 1..93
FT CDS /*tag= a
FT 94..1509 /*tag= b
FT /*product= "MAP kinase phosphatase splice form 1"
FT variation replace(577,A)
FT /*tag= c
FT /standard_name= "single nucleotide polymorphism (SNP)"
FT /*tag= d
FT /standard_name= "single nucleotide polymorphism (SNP)"
FT 3'UTR 1510..2704
FT variation /*tag= e
FT /*tag= (2641,A)
FT /*tag= f
FT /*standard_name= "single nucleotide polymorphism (SNP)"
XX
XX WO200242436-A2.
XX
XX 30-MAY-2002.
XX
XX 07-NOV-2001; 2001WO-US42995.
XX
XX 20-NOV-2000; 2000US-0715177.
XX 18-JAN-2001; 2001US-0761640.
XX
XX (PEKE ) PE CORP NY.
XX
XX WPI: 2002-575237/61.
XX P-PSDB: ABP51653.
XX
XX Novel isolated human phosphatase peptide useful for treating disorder
XX characterized by absence of, inappropriate or unwanted expression of
XX the phosphatase protein, and as immunogens to raise antibodies
XX
XX Claim 1; Fig 3A; 85bp; English.
XX
XX The present invention describes an isolated human phosphatase peptide
XX (I). (I) can be used for identifying a modulator of (I) by contacting
XX (I) with an agent and determining if the agent has modulated the
XX function or activity of (I). (I) is useful for identifying an agent that
XX binds to (I), by contacting (I) with an agent and assaying the contacted
XX mixture to determine whether a complex is formed with the agent bound
XX (I). The human phosphatases from the present invention are mitogen
XX activated protein (MAP) kinase phosphatases. These human MAP kinase
XX phosphatases are located on chromosome 11. (I) and the polynucleotide
XX sequences encoding (I) can be used in gene therapy. The present sequence
XX encodes human MAP kinase phosphatase splice form 1 from the present
XX invention.
XX
XX Sequence 2704 BP; 569 A; 874 C; 794 G; 467 T; 0 other;
XX
Query Match 100.0%; Score 2704; DB 24; Length 2704;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2704; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY	1	CGCCTTCCTGCTCTCCGGGGTCCAGACGTCGCCGGGGTTGAGGAAGGGCCGCTACC	60
Db	1	CGTCTTCCTGCTCTCCGGGGTCCAGACTGTCGGGGGTTGAGGAAGGGGGCCGTACC	60
QY	61	CGGTGCCAGCCAGGTCCTCGCGGCGCTGGCTCCATGGCCCTGGTCACAGTAGCCGTTCC	120
Db	61	CGGTGCCAGCCAGGTCCTCGCGGCGCTGGCTCCATGGCCCTGGTCACAGTAGCCGTTCC	120
QY	121	CCCCGGGACAGGGCGCTCTCCAGCCCTGAGGGCCCTGGGACCAAGCGGTCCAGCGAAG	180
Db	121	CCCCGGGACAGGGCGCTCTCCAGCCCTGAGGGCCCTGGGACCAAGCGGTCCAGCGAAG	180
QY	181	AGTCGACTCCAGCAAGGACAGAGCTTTGCGGTCTCCGTGGGGCTGTCTGGAGTCCAG	240
Db	181	AGTGCATCCCAAGGACAGAGAGCTTTGCGGTCTCCGTGGGGCTGTCTGGAGTCCAG	240
QY	241	GATGAGGGGGCAATGATGATGACAGAGAGCCAGTTCTTGACCAACAGAGAGCCCG	300
Db	241	GATGAGGGGGCAATGATGATGACAGAGAGCCAGTTCTTGACCAACAGAGAGCCCG	300
QY	301	AGTGAGGAGAGCTCCACAGGGGACAGACAGATTGGGGCAAGATGCCAGAGTCCCG	360
Db	301	AGTGAGGAGAGAGCTCCACAGGGGACAGACAGATTGGGGCAAGATGCCAGAGTCCCG	360
QY	361	AAGCAGAGAGAGCAGAGGACGACACTTCACCTCATGTAAGTCTCTGAGGCCGAGAT	420
Db	361	AAGCAGGAGAGCAGAGAGGACGACACTTCACCTCATGTAAGTCTCTGAGGCCGAGAT	420
QY	421	GACATCCGCGCGGAGGCCAGAGCTGGAGGACCCCGGCTCCCGGCTCGGACTGCTG	480
Db	421	GACATCCGCGCTGGAGGCCAGAGCTGGAGGACCCCGGCTCCCGGCTCGGACTGCTG	480
QY	481	GTAGTTTCTACAGAGAGAGGAAGGTCTGAGGCAGAGATGAGACGGTCTCTTGCGCTG	540
Db	481	GTAGTTTCTACAGAGAGGAAGGTCTGAGGCAGAGATGAGACGGTCTCTTGCGCTG	540
QY	541	GATTTCCCTGACACACTCTCCCGAGCTGACCTTGAGGCTGGCTGTGTGCTCTGAGGT	600
Db	541	GATTTCCCTGACACACTCTCCCGAGCTGACCTTGAGGCTGGCTGTGTGCTCTGAGGT	600
QY	601	GACACCCAGGTGTAAGTATGATGAGAGAGGGGGCTTGAGGTACGTCGTGGAGCAAGC	660
Db	601	GACACCCAGGTGTAAGTATGATGAGAGAGGGGGCTTGAGGTACGTCGTGGAGCAAGC	660
QY	661	CGGATCTTCAAGCCCATCTCCATCCAAACCATTTGGGCCACACTCAGGATATTGCACNA	720
Db	661	CGGATCTTCAAGCCCATCTCCATCCAAACCATTTGGGCCACACTCAGGATATTGCACNA	720
QY	721	GCATGTGAGGAGGCTTAAGGAGGAGGGGCTTGTAACGGGTGGCAGTGGCCCTCACGCGGCC	780
Db	721	GCATGTGAGGAGGCTTAAGGAGGAGGGGCTTGTAACGGGTGGCAGTGGCCCTCACGCGGCC	780
QY	781	AGCCACTACAGGAGAGACTGAATCCGGAACAGAGGTGCTCAATGATGAGAGCGGCTATG	840
Db	781	AGCCACTACAGGAGAGACTGAATCCGGAACAGAGGTGCTCAATGATGAGAGCGGCTATG	840
QY	841	GCCGACCTGGAGTCTCTGCGGCTCCAGAGGCGGACGCTGGGGGCTCTCAGAACAGAG	900
Db	841	GCCGACCTGGAGTCTCTGCGGCTCCAGAGGCGGACGCTGGGGGCTCTCAGAACAGAG	900
QY	901	CAGATGAGAGAGGAGATCCGGCTGAGCTGTGGAAGTTTGATATCTAGTACCTGGAG	960
Db	901	CAGATGAGAGAGGAGATCCGGCTGAGCTGTGGAAGTTTGATATCTAGTACCTGGAG	960
QY	961	AGTGTCACTTCCAAAGATCCGACAGGCTGTGAGAGTGCGGCTGGGGCTCCCTCCAG	1020
Db	961	AGTGTCACTTCCAAAGATCCGACAGGCTGTGAGAGTGCGGCTGGGGCTCCCTCCAG	1020
QY	1021	CAGTACCGTGACTTATCGAACAACAGATCTCTCTGTGGTGCACAGCGGAGCGAGCC	1080
Db	1021	CAGTACCGTGACTTATCGAACAACAGATCTCTCTGTGGTGCACAGCGGAGCGAGCC	1080

QY	1081	TCGGCATCTTCCCCACACCTCTACCTCGGGGCTCAGATGGAAACGACGACAAACCTGAGAG	1140
Db	1081	TCGGCATCTTCCCCACACCTCTACCTCGGGGCTCAGATGGAAACGACGACAAACCTGAGAG	1140
QY	1141	CTGCAGAGGAACAGGGTCACCCACATCTTGAACATAGGCCGGGAGATTGACAACTTCTAC	1200
Db	1141	CTGCAGAGGAACAGGGTCACCCACATCTTGAACATAGGCCGGGAGATTGACAACTTCTAC	1200
QY	1201	CCTGAGGGCTTCACTACCTACCAATGTGGGCTCTGGGATGAGAGACTGGGGCAGCTGCTG	1260
Db	1201	CCTGAGGGCTTCACTACCTACCAATGTGGGCTCTGGGATGAGAGACTGGGGCAGCTGCTG	1260
QY	1261	CCGCACTGGAAGAGAGACGCCCTTCAATTGAGGGCTGCAAGAGCACAGAGGCACCCACGTG	1320
Db	1261	CCGCACTGGAAGAGAGACGCCCTTCAATTGAGGGCTGCAAGAGCACAGAGGCACCCACGTG	1320
QY	1321	CTGGTCCACTGCAAGATGGGCTGACGCCGCTCAGCGGCTCACAGTGTGCTGTATGCCATG	1380
Db	1321	CTGGTCCACTGCAAGATGGGCTGACGCCGCTCAGCGGCTCACAGTGTGCTGTATGCCATG	1380
QY	1381	AAGCAGTACGAATCACCCTGAGACAGGCCCTGCGGCACAGTGCAGAGACTCGGCCCATC	1440
Db	1381	AAGCAGTACGAATCACCCTGAGACAGGCCCTGCGGCACAGTGCAGAGACTCGGCCCATC	1440
QY	1441	GCCCCCCCCAACCCTGGCTTCTCGGCGACGCTCAGATCTACAGGGCATCCTGACGGCC	1500
Db	1441	GCCCCCCCCAACCCTGGCTTCTCGGCGCGACGCTCAGATCTACAGGGCATCCTGACGGCC	1500
QY	1501	AGAACCTGAGGGTGTGGGAGAGAAAGGTTGTAGCATGGAAGAGACCCAGCAGGCC	1560
Db	1501	AGAACCTGAGGGTGTGGGAGAGAAAGGTTGTAGCATGGAAGAGACCCAGCAGGCC	1560
QY	1561	GAAGAAGAGAGCTGGGGCCACGGCCACGTATAAACCTCCGAGGGGTATAGGTCCTCATG	1620
Db	1561	GAAGAAGAGAGCTGGGGCCACGGCCACGTATAAACCTCCGAGGGGTATAGGTCCTCATG	1620
QY	1621	TCTTCTGGAGGCTTCCTTGGAGGCTGGAAGAGACCTCAGAGACAGATGAGCCAGAGGT	1680
Db	1621	TCTTCTGGAGGCTTCCTTGGAGGCTGGAAGAGACCTCAGAGACAGATGAGCCAGAGGT	1680
QY	1681	CTTCTCTTCCACAGAGCTTTCACATGAAGAGCCTCTGCAGCCCTTCCACAGCTTGCAG	1740
Db	1681	CTTCTCTTCCACAGAGCTTTCACATGAAGAGCCTCTGCAGCCCTTCCACAGCTTGCAG	1740
QY	1741	GACCAAGGAGGCCAGCAGAGTGGACAGAGGGGCTCTACGCTGCCCTGAAGTCCCCCACTC	1800
Db	1741	GACCAAGGAGGCCAGCAGAGTGGAGAGGGGGCTCTACGCTGCCCTGAAGTCCCCCACTC	1800
QY	1801	AGTGGTATACCTTCACAGGCGAGTCCCTGTGGGCGCAACCGGACCCAGGCTTCCAGAGCA	1860
Db	1801	AGTGGTATACCTTCACAGGCGAGTCCCGTGGTGGGCGCAACCGGACCCAGGCTTCCAGAGCA	1860
QY	1861	GGAGCAGGGGCGAGGGGCGAGGAGAGCCCTGCAATTTCTCTACGCCAGGTTCCG	1920
Db	1861	GGAGCAGGGGCGAGGGGCGAGGAGAGCCCTGCAATTTCTCTACGCCAGGTTCCG	1920
QY	1921	GAAGTGGTGAAGAGGCGACAGCGTCATGACAGTGGAGAGGGGCGAGGGCTGAGCCCT	1980
Db	1921	GAAGTGGTGAAGAGGCGACAGCGTCATGACAGTGGAGAGGGGCGAGGGCTGAGCCCT	1980
QY	1981	CACACATGCCCAACGCTCCCTGTGACACTGAAGAGATCCCAACACTCCTTGGAGAAAACACC	2040
Db	1981	CACACATGCCCAACGCTCCCTGTGACACTGAAGAGATCCCAACACTCCTTGGAGAAAACACC	2040
QY	2041	TCACGTCTGTGGCGGCACACATTCCTCTACGCTCGGCCCATACCGGTACTACAGGCTC	2100
Db	2041	TCACGTCTGTGGCGGCACACATTCCTCTACGCTCGGCCCATACCGGTACTACAGGCTC	2100
QY	2101	ACGTCGCCACCCCTGTCACTACGGGCTTACCTCCCAACCCCTGTCACTACAGCTCACTCC	2160
Db	2101	ACGTCGCCACCCCTGTCACTACGGGCTTACCTCCCAACCCCTGTCACTACAGCTCACTCC	2160
QY	2161	TACACGCTTAAATGCCAGGCCCATGTCTGTGCTGTCCAAAGGGCTCAAGACTTCTAATGG	2220

Db	2181	TACAGCCTTAAGTGTCCAGGCGCCATGTGTGTGCTGTCCAAAGGCTCAAGACTTTCATCTGG	2220
Qy	2221	GATGTGGTAGAGGACCTGTAAGGTACCTTTGGGGGCAACAGCACCCTAGTTTCATTCTCAA	2280
Db	2221	GATGTGGTAGAGGACCTGTAAGGTACCTTTGGGGGCAACAGCACCCTAGTTTCATTCTCAA	2280
Qy	2281	CTTACGCCCTGCACATCTCACCCTGTGGACAGGAATGAAACAGAGCTTCCCGTCAAAAG	2340
Db	2281	CTTACGCCCTGCACATCTCACCCTGTGGACAGGAATGAAACAGAGCTTCCCGTCAAAAG	2340
Qy	2341	GGTCACGCGCTCCACACCCCGCCCTCCCTGTGCACCTCTGTCTCTCCAGTTCAATCTCT	2400
Db	2341	GGTCACGCGCTCCACACCCCGCCCTCCCTGTGCACCTCTGTCTCTCCAGTTCAATCTCT	2400
Qy	2401	GGAACACGCCAGGCGCAGGCAACCAATGTGGCCCCCAAGGCAAGGCAAGATCTTAGGCCCA	2460
Db	2401	GGAACACGCCAGGCGCAGGCAACCAATGTGGCCCCCAAGGCAAGGCAAGATCTTAGGCCCA	2460
Qy	2461	GCCCGGAGGAGCTGGAAGGGCTGGCAGATCGCTTCCCTATCCACCTCACCGGATCCAGG	2520
Db	2461	GCCCGGAGGAGCTGGAAGGGCTGGCAGATCGCTTCCCTATCCACCTCACCGGATCCAGG	2520
Qy	2521	TCCTTGTCTGTCTGCCACAGCTCTGTGACACACACGACGACGATCAGACAGGACACAGGCCA	2580
Db	2521	TCCTTGTCTGTCTGCCACAGCTCTGTGACACACACGACGACGATCAGACAGGACACAGGCCA	2580
Qy	2581	GAGTACATCTTCTTTTGTGCTTCTGTGGCTCTGGCTGCTACCTAGTTTCATAGCTTCTACA	2640
Db	2581	GAGTACATCTTCTTTTGTGCTTCTGTGGCTCTGGCTGCTACCTAGTTTCATAGCTTCTACA	2640
Qy	2641	GTATCTGGCTTTGTACTGAGAAATATAACACATTTTCATATAAAAAAAAAAAAAAAAAA	2700
Db	2641	GTATCTGGCTTTGTACTGAGAAATATAACACATTTTCATATAAAAAAAAAAAAAAAAAA	2700
Qy	2701	AAAA 2704	
Db	2701	AAAA 2704	
RESULT 3			
AAD36063			
ID	AAD36063 standard; cDNA; 2618 BP.		
XX			
AC	AAD36063;		
XX			
DT	09-AUG-2002 (first entry)		
XX			
DE	Murine dual-specificity phosphatase 15 (DSP-15) cDNA.		
XX			
KW	Murine; dual-specificity phosphatase 15; DSP15; antiallergic; cytosolic;		
KW	immunosuppressive; MAP; mitogen activated protein kinase; cancer; enzyme;		
KW	signal transduction; cell proliferation; Duchenne muscular dystrophy;		
KW	cell cycle abnormality; graft-versus-host disease; autoimmune disease;		
KW	metabolic disease; allergy; screening; gene; ss.		
XX			
OS	Mus musculus.		
XX			
XX	Key Location/Qualifiers		
FT	35..1450		
FT	CDS		
FT	/*tag= a		
FT	/product= "Murine DSP-15 protein"		
XX			
FN	WO200224740-A2.		
XX			
PD	28-MAR-2002.		
XX			
PF	19-SEP-2001; 2001MO-US29406.		
XX			
PR	19-SEP-2000; 2000US-233833P.		
PR	18-SEP-2001; 2001US-0935732.		
XX			
PA	(CEPT-) CEPTYR INC.		

XX Luche RM, Wei B;
PI
XX
XX WPI: 2002-394127/42.
DR
DR P-PSDB: AAE22733.
XX
XX
XX
XX
XX
XX
PS
PS Claim 56; Fig 4; 91pp; English.
XX
XX
XX The invention relates to a new isolated dual-specificity phosphatase 15
CC (DSP-15) polypeptide which retains the ability to dephosphorylate an
CC activated MAP (mitogen activated protein) kinase. DSPs are phosphatases
CC that dephosphorylate both phosphotyrosine and phosphothreonine/serine
CC residues. DSP-15 polypeptides may be used to identify agents that
CC modulate DSP-15 activity, where such agents may inhibit or enhance signal
CC transduction via a MAP-kinase cascade, leading to cell proliferation. DSP
CC polypeptides, modulating agents, and/or polynucleotides encoding the
CC polypeptides may be used to modulate DSP-15 activity in a patient, and to
CC ameliorate disorders such as Duchenne muscular dystrophy, cancer, graft-
CC versus-host disease, autoimmune diseases, allergies, metabolic diseases,
CC abnormal cell growth, abnormal cell proliferation and cell cycle
CC abnormalities. DSP-15 alternate form polypeptides are useful in screening
CC assays for modulators of enzyme activity and/or substrate binding. The
CC present sequence is murine DSP-15 cDNA.
XX
XX Sequence 2618 BP; 538 A; 857 C; 769 G; 454 T; 0 other;
SQ

Query Match Similarity		96.8%;	Score 2618;	DB 24;	Length 2618;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 2618;	Conservative	0;	Mismatches	0;	Indels
					Gaps
					0;
OY	60	CCGGTCCACGCCAGGTGCTCGCGGCTTGCTGCATATGGCCCTGGTACAGTGAAGCCGTTTC	119		
Db	1	CCGGTGCACGCCAGCGATGCTCGCGGCTTGCTGCATATGGCCCTGGTACAGTGAAGCCGTTTC	60		
OY	120	GCCCGCCGGGACAGGCGGCTCCAGCCCGTGGGGGCGCTGGGACAGGGGGTCCAGCGAAG	179		
Db	61	GCCCCCGGGACAGGCGGCTCCAGCCCGTGGGGGCGCTGGGACAGGGGGTCCAGCGAAG	120		
OY	180	GAGTGACTCCACGACAAGGACAGAGCTTTGGGGTGCCTCGTGGGCGCTGTCCTGGACTCA	239		
Db	121	GAGTGACTCCACGACAAGGACAGAGCTTTGGGGTGCCTCGTGGGCGCTGTCCTGGACTCA	180		
OY	240	GGATGGAGGGGACATGATGATGCACGACAGAGGCCAGTTCTTGAGCCAAAGAGAGGCCCC	299		
Db	181	GGATGGAGGGGACATGATGATGCACGACAGAGGCCAGTTCTTGAGCCAAAGAGAGGCCCC	240		
OY	300	GAGTGAGGAGAGCTCCACGCGGGACAGACACACTTCGGGCAAGGATCCAGACTGCCCA	359		
Db	241	GAGTGAGGAGAGCTCCACGCGGGACAGACACACTTCGGGCAAGGATCCAGAGTCCCA	300		
OY	360	GAGGACGAGAGACAGAGGACAGCCTGCACCTCATGTGACAGTGTGAGGCGGCAGGA	419		
Db	301	GAGGACGAGGAGACAGAGGACAGCCTGCACCTCATGTGACAGTGTGAGGCGGCAGGA	360		
OY	420	TGACATCGCGCTGGAGGCCAGCTGGAGGACCCGGGCGTCCCGGGCTCCGCTACTCTGT	479		
Db	361	TGACATCGCGCTGGAGGCCAGCTGGAGGACCCGGGCGTCCCGGGCTCCGCTACTCTGT	420		
OY	480	GGTAGTTCTTACACAGAGAAGAGGTGTGAGCCAGGATGAGACGATCTTCTGGGGGT	539		
Db	421	GGTAGTTCTTCTACACAGAGAAGAGGTGTGAGCCAGGATGAGAGCGTCTTCTGGGGGT	480		
OY	540	GGATTTCCTGACACACACTCCCCAGCTGCACCTTGGGGCTGGTCTTGGGCTTGAG	599		
Db	481	GGATTTCCTGACACACACTCCCCAGCTGCACCTTGGGGCTGGTCTTGGGCTTGAG	540		
OY	600	TGACACCCAGAGTGTCTTAGATGAGAGCGGGGGCTTCAGCTGACGTCGTGGTGGCAAG	659		

Db 541 TGACACCCAGGTGTACTAGATGAGACGGGGGCTTCAAGCTGACGTGCTGGTGGCCAAAG 600
 QY 660 CCGATCTTCAAGCCCATCTCCATCCAGACCATGTGGGCCAAGCTCCAGATATTCACCA 719
 Db 601 CCGATCTTCAAGCCCATCTCCATCCAGACCATGTGGGCCAAGCTCCAGATATTCACCA 660
 QY 720 AGCATGTGAGGAGCTCTAGAGCAGCGGCTTGTACCGGGTGGCACTGCCCTACCTGGGC 779
 Db 661 AGCATGTGAGGAGCTCTAGAGCAGCGGCTTGTACCGGGTGGCACTGCCCTACCTGGGC 720
 QY 780 CAGCCACTCCAGAGAGAGCTGAACCCGAAACAGAGCTCCCTCAATGATGTGAGCGGCTAT 839
 Db 721 CAGCCACTCCAGAGAGAGCTGAACCCGAAACAGAGCTCCCTCAATGATGTGAGCGGCTAT 780
 QY 840 GGCAGCAGCTGAGTCTGTGGGCTCCAGCCGCGAGCCTGGCGGGTCTCTAGAAACAGA 899
 Db 781 GGCAGCAGCTGAGTCTGTGGGCTCCAGCCGCGAGCCTGGCGGGTCTCTAGAAACAGA 840
 QY 900 GCAGATGAGAGAGGAGTCCGTGTGAGTGTGAAAGTGTGATGTGATGAGTACCTGGA 959
 Db 841 GCAGATGAGAGAGGAGTCCGTGTGAGTGTGAAAGTGTGATGTGATGAGTACCTGGA 900
 QY 960 GAGTGTCACTTCCAAAGAGATCCGCGAGGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1019
 Db 901 GAGTGTCACTTCCAAAGAGATCCGCGAGGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 960
 QY 1020 GCAGTACCTGATCTTCAATGAGACCAACAGATGCTGTGTGTGGCAGCAGCGGAGCCAGC 1079
 Db 961 GCAGTACCTGATCTTCAATGAGACCAACAGATGCTGTGTGTGGCAGCAGCGGAGCCAGC 1020
 QY 1080 CTCGCCATCTTCCCGACCTCTACCTGGGCTCAGAGTGGAAAGCGCAAACTGGAGGA 1139
 Db 1021 CTCGCCATCTTCCCGACCTCTACCTGGGCTCAGAGTGGAAAGCGCAAACTGGAGGA 1080
 QY 1140 GCTGCAAGAGAACAGGGGTACCCACATCTTGAACATGAGCCCGGAGATTCACAACTTCTA 1199
 Db 1081 GCTGCAAGAGAACAGGGGTACCCACATCTTGAACATGAGCCCGGAGATTCACAACTTCTA 1140
 QY 1200 CCTGTAGCCCTTACCTACCAATGTGGGCTCTGGGATGAGAGTGGAGCCAGCTGCT 1259
 Db 1141 CCTGTAGCCCTTACCTACCAATGTGGGCTCTGGGATGAGAGTGGAGCCAGCTGCT 1200
 QY 1260 GCGGCACTGGAGAGAACGACCGCTTCAATGAGCTGCAAGAGCAAGGGGACCCACGT 1319
 Db 1201 GCGGCACTGGAGAGAACGACCGCTTCAATGAGCTGCAAGAGCAAGGGGACCCACGT 1260
 QY 1320 GCTGTGTCACTGCAAGATGGGCTGAGCCGCTGAGCGGACAGTCTGTGGCTATGCCAT 1379
 Db 1261 GCTGTGTCACTGCAAGATGGGCTGAGCCGCTGAGCGGACAGTCTGTGGCTATGCCAT 1320
 QY 1380 GAAGCAGTACGAATGAGCCTGTGAGCAGGCGCTGCGCCACAGTGCAGAGCTCGGCGCAT 1439
 Db 1321 GAAGCAGTACGAATGAGCCTGTGAGCAGGCGCTGCGCCACAGTGCAGAGCTCGGCGCAT 1380
 QY 1440 CGCGCGCCCAACCTGTGGCTTCTGTGGCCAGCTGCAAGATCTACAGGGGCTCTGTAGGCG 1499
 Db 1381 CGCGCGCCCAACCTGTGGCTTCTGTGGCCAGCTGCAAGATCTACAGGGGCTCTGTAGGCG 1440
 QY 1500 CAAACCTGAGGCTGTGGGAGAGAGTGTAGGCATGAAGAGACAGCAGGAGCC 1559
 Db 1441 CAAACCTGAGGCTGTGGGAGAGAGTGTAGGCATGAAGAGACAGCAGGAGCC 1500
 QY 1560 CGAAGAGAGCCTGGGCGCAGGCGCATATAAAGCTCCAGAGGGGTCTAGAGTGCATCA 1619
 Db 1501 CGAAGAGAGCCTGGGCGCAGGCGCATATAAAGCTCCAGAGGGGTCTAGAGTGCATCA 1560
 QY 1620 GTCTTCTGAGCCCTCTTGGAGCTGAGAGCAGCTGAGAGCAGTGAATGCCAGAGG 1679
 Db 1561 GTCTTCTGAGCCCTCTTGGAGCTGAGAGCAGCTGAGAGCAGTGAATGCCAGAGG 1620
 QY 1680 TCTTCTTCCAGAGTCTTCAATGAAGAGCCTGAGAGCCCTCCAGAGTTGCA 1739
 Db 1621 TCTTCTTCCAGAGTCTTCAATGAAGAGCCTGAGAGCCCTCCAGAGTTGCA 1680

QY 1740 GGACCAAGGAGCCAGCAGTGTGACAGAGGGGCTTCAAGCTGCTGCTTGAAGTCCCGCAGT 1799
 Db 1681 GGACCAAGGAGCCAGCAGTGTGACAGAGGGGCTTCAAGCTGCTGCTTGAAGTCCCGCAGT 1740
 QY 1800 CAGTGTATACCTCCAGGCGAGTGGCGGCTGTGGGCAACCGGAGCCAGGCTTCCAGAGC 1859
 Db 1741 CAGTGTATACCTCCAGGCGAGTGGCGGCTGTGGGCAACCGGAGCCAGGCTTCCAGAGC 1800
 QY 1860 AGAGCAGAGGCGAGGCGCAGGAGAGAGCCCTGATTTCTCTACGCCAGGTTCC 1919
 Db 1801 AGAGCAGAGGCGAGGCGCAGGAGAGAGCCCTGATTTCTCTACGCCAGGTTCC 1860
 QY 1920 GGAAGTGTGTAGACAGGCGCAGCGTGCATGACAGTGGAGAGAGGAGGCGGCTGAGCC 1979
 Db 1861 GGAAGTGTGTAGACAGGCGCAGCGTGCATGACAGTGGAGAGAGGAGGCGGCTGAGCC 1920
 QY 1980 TCACACATGCGCAGCTCCCTGACACATGAAGAGATCCAAAGCTCTTGGAGAAGAC 2039
 Db 1921 TCACACATGCGCAGCTCCCTGACACATGAAGAGATCCAAAGCTCTTGGAGAAGAC 1980
 QY 2040 CTCACGTGTGTTCCGACACATTTCTCTCAGCTCCGCCCATACCGGTCACTACAGCT 2099
 Db 1981 CTCACGTGTGTTCCGACACATTTCTCTCAGCTCCGCCCATACCGGTCACTACAGCT 2040
 QY 2100 CACCTCCCAACCCCTGTCACTACAGGCTCAGCTCCGACCCCTGTCACTACAGCTCAGCTC 2159
 Db 2041 CACCTCCCAACCCCTGTCACTACAGGCTCAGCTCCGACCCCTGTCACTACAGCTCAGCTC 2100
 QY 2160 CTACAGCTTAAAGTCCCAAGGCGCATGTGCTGTCTGCAAGGAGCTCAAGACTTCTAACTG 2219
 Db 2101 CTACAGCTTAAAGTCCCAAGGCGCATGTGCTGTCTGCAAGGAGCTCAAGACTTCTAACTG 2160
 QY 2220 GGATGTGTAGAGGAGCAGTGAAGTACCTTTGGGGCAACAGCACCCTAGTTCAATCTCA 2279
 Db 2161 GGATGTGTAGAGGAGCAGTGAAGTACCTTTGGGGCAACAGCACCCTAGTTCAATCTCA 2220
 QY 2280 ACTTAAGCCCTGACACACCTGAGTGGCAGGAGTGAAGAAAGAGACTTCCGTGCAAAA 2339
 Db 2221 ACTTAAGCCCTGACACACCTGAGTGGCAGGAGTGAAGAAAGAGACTTCCGTGCAAAA 2280
 QY 2340 GGGTCAAGCCTCCACCCCGGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2399
 Db 2281 GGGTCAAGCCTCCACCCCGGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2340
 QY 2400 TGAACACAGCAGGCGCAGGCAACAGTGGCCGCCAAAGGCAAGGAGTCTCAGGCCCC 2459
 Db 2341 TGAACACAGCAGGCGCAGGCAACAGTGGCCGCCCAAAGGCAAGGAGTCTCAGGCCCC 2400
 QY 2460 AGCGGGGAGGCTGGAAGGCTGGCAGATGGCTTCCGATCCAGCTCCAGCGGTCAG 2519
 Db 2401 AGCGGGGAGGCTGGAAGGCTGGCAGATGGCTTCCGATCCAGCTCCAGCGGTCAG 2460
 QY 2520 GTCTTGTGCTGTGCTCCAGACCTCTGTGACACACAGCAGCAATCAGAGGGACAGGCG 2579
 Db 2461 GTCTTGTGCTGTGCTCCAGACCTCTGTGACACACAGCAGCAATCAGAGGGACAGGCG 2520
 QY 2580 AGAGATAGTCTTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2639
 Db 2521 AGAGATAGTCTTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2580
 QY 2640 AGTATGTGCTTGTGCTGAGAAATTAACACATTTTC 2677
 Db 2581 AGTATGTGCTTGTGCTGAGAAATTAACACATTTTC 2618

RESULT 4
 AAF30485
 ID AAF30485 standard; cDNA: 2781 BP.
 XX AAF30485;
 AC
 XX
 DT 29-MAY-2001 (first entry)

QY 1157 TCACCCACATCTTGAACATGGCCCGGAGATTGACAACTTCTACCTGAGCGCTTACCT 1216
 Db 1150 TCACCCACATCTTGAACATGGCCCGGAGATTGACAACTTCTACCTGAGCGCTTACCT 1209
 QY 1217 ACCACAATGTCGGCTCTGGGATGAGAGATCGGCCAGCTGCTGCCGACCTGGAAGGAGA 1276
 Db 1210 ACCACAATGTCGGCTCTGGGATGAGAGATCGGCCAGCTGCTGCCGACCTGGAAGGAGA 1269
 QY 1277 GCACACGCTTCTTGAAGTGCAGAGACAGAGGACCCAGCTGCTGATGCTCAGTCAAGA 1336
 Db 1270 GCACACGCTTCTTGAAGTGCAGAGACAGAGGACCCAGCTGCTGATGCTCAGTCAAGA 1329
 QY 1337 TGGGCGCTCAGCCGCTCAGCGGCCACAGTGTGGCTATGSCATGAAGAGTAACTGCA 1396
 Db 1330 TGGGCGCTCAGCCGCTCAGCGGCCACAGTGTGGCTATGSCATGAAGAGTAACTGCA 1389
 QY 1397 GCCTGAGAGAGGCTTGGCCGACAGTGCAGAGAGTCTGGGCGCATGCGCCGCCCAACCTTG 1456
 Db 1390 GCCTGAGAGAGGCTTGGCCGACAGTGCAGAGAGTCTGGGCGCATGCGCCGCCCAACCTTG 1449
 QY 1457 GCTTCTGGCCAGCTGCAGATCTACAGGGCATCTGACG----- 1497
 Db 1450 GCTTCTGGCCAGCTGCAGATCTACAGGGCATCTGACAGGGCATCTGACAGGGCATG 1509
 QY 1498 ----- 1497
 Db 1510 TCTGGAGAGAGAAATGGGTGGGTCTCCCAAGAGAGCACCCAGCCCTGAAGTCTTA 1569
 QY 1498 -----GCCAAGACCTGAGGGTGGTGGGAGGAGGAGTGTAGGCA 1538
 Db 1570 CACCATTTCCACCTCTTCGCGCAGAACTGAGGGTGTGGGAGGAGAGTGTAGGCA 1629
 QY 1539 TGGAGAAGAGCCAGGAGGCCCGGAAAGAGAGCTGGGCCAGCGCATTAATACTCC 1598
 Db 1630 TGGAGAAGAGCCAGGAGGCCCGGAAAGAGAGCTGGGCCAGCGCATTAATACTCC 1689
 QY 1599 GAGGGGTGATGAGGTCCATCATGTTCTGTGAGAGCCCTCTTGGAGCTGGAGAGCACTCAG 1658
 Db 1690 GAGGGGTGATGAGGTCCATCATGTTCTGTGAGAGCCCTCTTGGAGCTGGAGAGCACTCAG 1749
 QY 1659 AGACCAAGTGCATGTCAGAGAGTCTTCCTCCAGAGAGTTCACATGAAGAGCTCTGC 1718
 Db 1750 AGACCAAGTGCATGTCAGAGAGTCTTCCTTCCTCCAGAGTTCACATGAAGAGCTCTGC 1809
 QY 1719 AGCCCTTCCACAGCTTGAAGAGCAAGGAGGCCAGAGTGTGAACAGGGGCTCAGC 1778
 Db 1810 AGCCCTTCCACAGCTTGAAGAGCAAGGAGGCCAGAGTGTGAACAGGGGCTCAGC 1869
 QY 1779 CTGACCTGAAGTCCCGCAGTCAAGTGTACCTCCAGGGCAGTGGCTGTGGCCAAAC 1838
 Db 1870 CTGACCTGAAGTCCCGCAGTCAAGTGTACCTCCAGGGCAGTGGCTGTGGCCAAAC 1929
 QY 1839 GGAACCCAGGCTTCCAGAGACAGAGAGGGGAGGGGAGGGGAGGGAGAGCCCTGCA 1898
 Db 1930 GGAACCCAGGCTTCCAGAGACAGAGAGGGGAGGGGAGGGGAGGGAGAGCCCTGCA 1989
 QY 1899 TTTCTCTACAGCCAGGTTCGGAAGGTGTGAGACAGGCCAGCGTGCATGACAGTGGAG 1958
 Db 1990 TTTCTCTACAGCCAGGTTCGGAAGGTGTGAGACAGGCCAGCGTGCATGACAGTGGAG 2049
 QY 1959 AGGAGGCGAGGCTGAGCCCTCACACATGCCCCAGCTCCCTGACACTGAAGAGATCC 2018
 Db 2050 AGGAGGCGAGGCTGAGCCCTCACACATGCCCCAGCTCCCTGACACTGAAGAGATCC 2109
 QY 2019 ACAACTCTCTGGAGAAACACACCTCAGCTGTGTGCGGACACAACTCTCTCAGTCCGCC 2078
 Db 2110 ACAACTCTCTGGAGAAACACACCTCAGCTGTGTGCGGACACAACTCTCTCAGTCCGCC 2169
 QY 2079 CCATACCCGCTACTAGAGCTCAGCTCCACACCTGTCTACTAGAGGCTCTACCTCCACC 2138
 Db 2170 CCATACCCGCTACTAGAGCTCAGCTCCACACCTGTCTACTAGAGGCTCTACCTCCACC 2229

QY 2139 CTGTCTACAGCTCAGCTCAGCTTACAGCTTAAGTCCAGGCGCCATGTCCTTCCAA 2198
 Db 2230 CTGTCTACAGCTCAGCTCAGCTTACAGCTTAAGTCCAGGCGCCATGTCCTTCCAA 2289
 QY 2199 GGGCTCAGAGCTTCTTAAGTGTGTGAGAGGAGTGAAGTACCTTTGGGGGCAAC 2258
 Db 2290 GGGCTCAGAGCTTCTTAAGTGTGTGAGAGGAGTGAAGTACCTTTGGGGGCAAC 2349
 QY 2259 AGCACCCTAATTTCTATCTCAACTCTAGCCCTGCACACTCAGTGTGGACGGATGAA 2318
 Db 2350 AGCACCCTAATTTCTATCTCAACTCTAGCCCTGCACACTCAGTGTGGACGGATGAA 2409
 QY 2319 ACAGAGCTTCCCGTCAAAAAGGGTCAAGCTCCACCCCGCCCGCTCCCTGACCTCC 2378
 Db 2410 ACAGAGCTTCCCGTCAAAAAGGGTCAAGCTCCACCCCGCCCGCTCCCTGACCTCC 2469
 QY 2379 TGTCTCTCCAGTTCAATTCCTGTAACAGCAGCAGGACAGGACAGTGGCCCCAAAG 2438
 Db 2470 TGTCTCTCCAGTTCAATTCCTGTAACAGCAGCAGGACAGGACAGTGGCCCCAAAG 2529
 QY 2439 CAGGAGAGTCTCTAGGCCCGCAGCCGCGGAGGCTGGAAGGCTGGCAGATCGCTTCCCT 2498
 Db 2530 CAGGAGAGTCTCTAGGCCCGCAGCCGCGGAGGCTGGAAGGCTGGCAGATCGCTTCCCT 2589
 QY 2499 CATCCACCTCCACGGGCTCAGGCTTTGTGCTGCTGCCAGACCTCTGTGACACACAGC 2558
 Db 2590 CATCCACCTCCACGGGCTCAGGCTTTGTGCTGCTGCCAGACCTCTGTGACACACAGC 2649
 QY 2559 CAGATCACAGGGCAGCAGGCGCAGAGATAGTCTTCTTTGTCTCTTGTGCTGAGCTA 2618
 Db 2650 CAGATCACAGGGCAGCAGGCGCAGAGATAGTCTTCTTTGTCTCTTGTGCTGAGCTA 2709
 QY 2619 GTTCAGTTTTCATATGCTTACAGTATCTGTGCTTTGTACTGAGAAATAACACATTTC 2678
 Db 2710 GTTCAGTTTTCATATGCTTACAGTATCTGTGCTTTGTACTGAGAAATAACACATTTC 2769
 QY 2679 TAAAAAATAA 2690
 Db 2770 TAAAAAATAA 2781
 RESULT 5
 AB073250
 ID AB073250 standard: cDNA; 2852 BP.
 XX AC AB073250;
 XX DT 30-SEP-2002 (first entry)
 DE Human MAP kinase phosphatase splice form 2 cDNA sequence SEQ ID NO:2.
 KW Human; phosphatase; mitogen activated protein kinase phosphatase;
 KW MAP kinase; enzyme; chromosome 11; gene; ss.
 OS Homo sapiens.
 XX key Location/Qualifiers
 FH 1..56
 FT 5'UTR
 FT CDS
 FT /*tag= a
 FT /*tag= b
 FT product= "MAP kinase phosphatase splice form 2"
 FT 2037..2852
 FT /*tag= c
 XX PN MO200242436-A2.
 XX PD 30-MAY-2002.
 XX PF 07-NOV-2001; 2001WO-US42995.
 XX PR 20-NOV-2000; 2000US-0715177.
 PR 18-JAN-2001; 2001US-0761640.

D	b	1682	TCTTCTTGAGACCCCTCCTTGAGAGCTGAGAGCAACCTCACAGAACAGTAGCATGCCAAGGT	1741
O	y	1681	CTTCTCTTTCCACGAGCTCTTCACATGABAGGCCTTGCAGGCCCTTCCACAGCTTGGAA	1740
D	b	1742	CTTCTCTTCCACGAGAGCTCTTCACATGABAGGCCTTGCAGGCCCTTCCACAGCTTGGAA	1801
O	y	1741	GACCAGAAGGAGCGACAGAGGTGACAGGGGGGCTCAGCCTGCCCTGAAGTCCCSCAGTC	1800
D	b	1802	GACCAAGAGGAGCGACAGAGGTGACAGGGGGGCTCAGCCTGCCCTGAAGTCCCSCAGTC	1861
O	y	1801	AGTGGTTACCCCTCCAGGGCAGTGCCTGTGTGGCCAACGGACCAAGGCTTCCAGAGCA	1860
D	b	1862	AGTGGTTACCCCTCCAGGGCAGTGCCTGTGTGGCCAACGGACCAAGGCTTCCAGAGCA	1921
O	y	1861	GGACAGAGGGGACAGGGGCGAGGGGAGAGGCCGTAATTCTCTACGGCAGATTCCG	1920
D	b	1922	GGAGAGAGGGGACAGGGGCGAGGGGAGAGGCCGTAATTCTCTACGGCAGATTCCG	1981
O	y	1921	GAAGGTGGTGNAGACAGGCCAGCGCTGCATGACAGTGGAGAGAGGAGGCGAGGCTGAGCCT	1980
D	b	1982	GAAGGTGGTGNAGACAGGCCAGCGCTGCATGACAGTGGAGAGAGGAGGCGAGGCTGAGCCT	2041
O	y	1981	CACACATGCCCCAGCCTCCCTGACACTGAAGAGATCCACAACCTCTTGGAGAAAACACC	2040
D	b	2042	CACACATGCCCCAGCCTCCCTGACACTGAAGAGATCCACAACCTCTTGGAGAAAACACC	2101
O	y	2041	TCACGTCTGTGGCGGACACATTCCTCTAGCTACGCGCCCATACCCGTCACATAAGCCTC	2100
D	b	2102	TCACGTCTGTGGCGGACACATTCCTCTAGCTACGCGCCCATACCCGTCACATAAGCCTC	2161
O	y	2101	ACCTCCCAACCCCTGTCACTAGAGGCGTACCTCCACCCCTGTCACTACAGCCTCACTCC	2160
D	b	2162	ACCTCCCAACCCCTGTCACTAGAGGCGTACCTCCACCCCTGTCACTACAGCCTCACTCC	2221
O	y	2161	TCACGCTTAAGTCCCAAGGCCCAATGTGCGCTTCCCAAGGGCTCAAGACTTTCTAACGG	2220
D	b	2222	TCACGCTTAAGTCCCAAGGCCCAATGTGCGCTTCCCAAGGGCTCAAGACTTTCTAACGG	2281
O	y	2221	GATGTGTTAGAGGAGACTGAAGTACCTTTTG6GGGCAACAGCAACCCGTAATTTCATTC	2280
D	b	2282	GATGTGTTAGAGGAGACTGAAGTACCTTTTG6GGGCAACAGCAACCCGTAATTTCATTC	2341
O	y	2281	CTCTAGCCCTGACACATCACCTGTGGGACAGGAATGAANAAGAGGCTCCCGTCAAAAAG	2340
D	b	2342	CTCTAGCCCTGACACATCACCTGTGGGACAGGAATGAANAAGAGGCTCCCGTCAAAAAG	2401
O	y	2341	GGTCAAGGCTCCCAACCCCGGCCCTCCTGTGACACCTCTGTCTCTCCAGTTCATTCCT	2400
D	b	2402	GGTCAAGGCTCCCAACCCCGGCCCTCCTGTGACACCTCTGTCTCTCCAGTTCATTCCT	2461
O	y	2401	GGAACACAGCGCAGGCGAGGCAACAGTGGCCCCCAAAGGACAGGAGATCTTAGAGCCCA	2460
D	b	2462	GGAACACAGCGCAGGCGAGGCAACAGTGGCCCCCAAAGGAGAGGAGATCTTAGAGCCCA	2521
O	y	2461	GGCGGGGAGGCTGGAAAGGCTGGGCAATGGCTTCCCTATCCACTCAGCCGATCCAGG	2520
D	b	2522	GGCGGGGAGGCTGGAAAGGCTGGGCAATGGCTTCCCTATCCACTCAGCCGATCCAGG	2581
O	y	2521	TCTTTGCTGCTGTCCCAAGACCTCTGTGACACACGACGCAATCACAGGGCACAGGCCA	2580
D	b	2582	TCTTTGCTGCTGTCCCAAGACCTCTGTGACACACGACGCAATCACAGGGCACAGGCCA	2641
O	y	2581	GAGATAGCTCTTTTTGTCTTCTGTCTGGCTCTGGCGTAGCACTTTTCTATAGCCCTTCA	2640
D	b	2642	GAGATAGCTCTTTTTGTCTTCTGTCTGGCTCTGGCGTAGCACTTTTCTATAGCCCTTCA	2701
O	y	2641	GATATCGGCTTGTACTAGAAATAAAAACATTTTCATAAAAAAGAAAAAAAAAAAAA	2700
D	b	2702	GATATCGGCTTGTACTAGAAATAAAAACATTTTCATAAAAAAGAAAAAAAAAAAAA	2761
O	y	2701	AAAA 2704 	

Db 2762 AAAA 2765

RESULT 6
AAD36061
ID AAD36061 standard; cDNA: 2718 BP.
XX
AC AAD36061;
XX
DT 09-AUG-2002 (first entry)
XX
DE Human dual-specificity phosphatase 15 (DSP-15) cDNA.
XX
KW Human; dual-specificity phosphatase 15; DSP-15; antiallergic; cytosolic;
KW immunosuppressive; MAP; mitogen activated protein kinase; cancer; enzyme;
KW signal transduction; cell proliferation; Duchenne muscular dystrophy;
KW cell cycle abnormality; graft-versus-host disease; autoimmune disease;
KW metabolic disease; allergy; screening; chromosome 11q; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FH CDS 35..2014
FH FT /*tag= a
FT FT /product= "Human DSP-15 protein"
FT
PN WO200224740-A2.
PN
PD 28-MAR-2002.
PD
PF 19-SEP-2001; 2001WO-US29406.
PF
XX 19-SEP-2000; 2000US-233833P.
XX PR 18-SEP-2001; 2001US-0955732.
XX
XX (CEPT-) CEPTYR INC.
XX
XX Luche RM, Wei B;
XX
XX WPI: 2002-394127/42.
XX DR P-PSDB: AAE22729.
XX
XX New dual-specificity phosphatase 15 polypeptide and polynucleotides,
XX useful for treating e.g. Duchenne muscular dystrophy, cancer,
XX graft-versus-host disease, autoimmune diseases, allergies, metabolic
XX diseases
XX
XX Claim 7; Fig 1; 91pp; English.
XX
XX The invention relates to a new isolated dual-specificity phosphatase 15
XX (DSP-15) polypeptide which retains the ability to dephosphorylate an
XX activated MAP (mitogen activated protein) kinase. DSPs are phosphatases
XX that dephosphorylate both phosphotyrosine and phosphothreonine/serine
XX residues. DSP-15 polypeptides may be used to identify agents that
XX modulate DSP-15 activity, where such agents may inhibit or enhance signal
XX transduction via a MAP-kinase cascade, leading to cell proliferation. DSP
XX polypeptides, modulating agents, and/or polynucleotides encoding the
XX polypeptides may be used to modulate DSP-15 activity in a patient, and to
XX ameliorate disorders such as Duchenne muscular dystrophy, cancer, graft-
XX versus-host disease, autoimmune diseases, allergies, metabolic diseases,
XX abnormal cell growth, abnormal cell proliferation and cell cycle
XX abnormalities. DSP-15 alternate form polypeptides are useful in screening
XX assays for modulators of enzyme activity and/or substrate binding. The
XX present sequence is human DSP-15 cDNA. Human DSP-15 gene is located on
XX chromosome 11q.
XX
XX Sequence 2718 BP; 559 A; 893 C; 795 G; 471 T; 0 other;
XX
XX Query Match 92.9%; Score 2512; DB 24; Length 2718;
XX Best Local Similarity 96.4%; Pred. No. 0;
XX Matches 2620; Conservative 0; Mismatches 0; Indels 98; Gaps 1

60 CCGGTGCAAGCCCAAGGTGCTCGGGGCTGCTGCATGCGCCCTGTGTCAGATGAGCGGTTT 119

|||||
Db 1 CCGGTCGACGCCAGGTGCTCGGGCTGCTGCATGGCCCTGCTCACAGTGAAGCGCTTC 60
OY 120 GCCCGCGGACAGCGGCGCTCCAGCGCCGTGGGGCCCTGGGACGACGGGGTCCAGGGAAG 179
Db 61 GCCCGCGGACAGCGGCGCTCCAGCGCCGTGGGGCCCTGGGACGACGGGGTCCAGGGAAG 120
OY 180 GAGTCGACTCCAGCGACGAGGACGAGCTTTGGGCTGCTCCGTGGGGGTGTTCTGGGACTGA 239
Db 121 GAGTGAAGGAGGCTCCAGCGGAGGACGAGCTTTGGGCTGCTCCGTGGGGGTGTTCTGGGACTGA 180
OY 240 GGATGGAGGGGCAATGATGATGACAGAGGCGCAATTCTGAGCCAAACAGAGAAGGCCCC 299
Db 181 GGATGGAGGGGCAATGATGATGACAGAGGCGCAATTCTGAGCCAAACAGAGAAGGCCCC 240
OY 300 GAGTGAAGGAGGCTCCAGCGGAGGACGACAGCTTCGGGGCAAGGATCCCAAGATCCCA 359
Db 241 GAGTGAAGGAGGCTCCAGCGGAGGACGACAGCTTCGGGGCAAGGATCCCAAGATCCCA 300
OY 360 GAAGCAGAGGAGCAGAGGACGACACCTGCACTCATGTATGACAGCTGCTGAGGCGCGAGGA 419
Db 301 GAAGCAGAGGAGGAGCAGAGGACGACACCTGCACTCATGTATGACAGCTGCTGAGGCGCGAGGA 360
OY 420 TCACATCCGCTGCGACGCCAGCTGAGAGCACCCGGGCTCCCGGCTCCGCTACCTGCT 479
Db 361 TCACATCCGCTGCGAGCGCCAGCTGAGAGCACCCGGGCTCCCGGCTCCGCTACCTGCT 420
OY 480 GGTAGTTCTTACAGGAGGAGGAGGAGTGTGAGCCAGATGAGAGGAGTCTCTGGGGGT 539
Db 421 GGTAGTTCTTACAGGAGGAGGAGGAGTGTGAGCCAGATGAGAGGAGTCTCTGGGGGT 480
OY 540 GGATTTCCCTGACAGCAGCTCCCGAGCTGCACCTGGGGCTGCTTTCCTCCCTCTGAG 599
Db 481 GGATTTCCCTGACAGAGCTCCCGAGCTGCACCTGGGGCTGCTTTCCTCCCTCTGAG 540
OY 600 TCACACCCAGGCTGTACTTATGATGAGAGCGGGGCTTCAGCGTGAAGTCTGGTGGCAAG 659
Db 541 TCACACCCAGGCTGTACTTATGATGAGAGCGGGGCTTCAGCGTGAAGTCTGGTGGCAAG 600
OY 660 CCGGATCTTCAAGCCCATTCATCCAGACGATGTTGGGGCACATCCAGATTTTGCACCA 719
Db 601 CCGGATCTTCAAGCCCATTCATCCAGACGATGTTGGGGCACATCCAGATTTTGCACCA 660
OY 720 AGCATGTGAGGAGCTGTATGAGCAGCGGCTTGTACCGGGTGGCAGTGGCTTCACTGGGC 779
Db 661 AGCATGTGAGGAGCTGTATGAGCAGCGGCTTGTACCGGGTGGCAGTGGCTTCACTGGGC 720
OY 780 CAGCCACTACACGAGAGAGCTGACTCCGACAGAGCTGCTCAATGATGAGAGGCTAT 839
Db 721 CAGCCACTACACGAGAGAGCTGACTCCGACAGAGCTGCTCAATGATGAGAGGCTAT 780
OY 840 GGCAGCCTGAGGCTGCTCGGGGCTCCCAAGGCGGAGGCTGGGGGTCTCTCAAGACAGA 899
Db 781 GGCAGCCTGAGGCTGCTCGGGGCTCCCAAGGCGGAGGCTGGGGGTCTCTCAAGACAGA 840
OY 900 GCAGATGAGAGCAGGAGATCCGTGTAGCTGTGAAGGTGTGATGTGATGAGTACCTGGA 959
Db 841 GCAGATGAGAGCAGGAGATCCGTGTAGCTGTGAAGGTGTGATGTGATGAGTACCTGGA 900
OY 960 GAGTGTCACTTCCAAAGAGATCCGCCAGGCTGTGAGCTGCGCTGGGGCTCCCTCTCA 1019
Db 901 GAGTGTCACTTCCAAAGAGATCCGCCAGGCTGTGAGCTGCGCTGGGGCTCCCTCTCA 960
OY 1020 GAGTACGCTGACTCATGAGACGACGATGCTGCTGCTGGTGGGACAGCGGAGCGAGAC 1079
Db 961 GAGTACGCTGACTCATGAGACGACGATGCTGCTGCTGGTGGGACAGCGGAGCGAGAC 1020
OY 1080 CTCGCCATCTTCCCGACCTCTACCTGGGCTCAGAGTGAAGCAGAGCAAACTGGAGA 1139
Db 1021 CTCGCCATCTTCCCGACCTCTACCTGGGCTCAGAGTGAAGCAGAGCAAACTGGAGA 1080
OY 1140 GCTGAGAGGAGGAGGCTACCCACATCTTGAACATGCGCGGAGATTTGACAACTTCTA 1199
|||||

Db 1081 GCTGAGAGGAGGAGGCTACCCACATCTTTGAACATGGCCCGGAGATTGACAACTTCTA 1140
OY 1200 CCTGAGGCTTACACTTACACATATGTGGCCCTTGGGATGAGAGGTGGCCAGCTGCT 1259
Db 1141 CCTGAGGCTTACACTTACACATATGTGGCCCTTGGGATGAGAGGTGGCCAGCTGCT 1200
OY 1260 GCCGCACTGGAGGAGACGACCGCTTCAATGAGGCTGCAAGAGCAGACAGGACCCAGCT 1319
Db 1201 GCCGCACTGGAGGAGGAGCAGCGCTTCAATGAGGCTGCAAGAGCAGACAGGACCCAGCT 1260
OY 1320 GCTGTTCACCTCAAGATGGGCTGACCGCTGACGGGCGACAGTGTGGCTATGGCAT 1379
Db 1261 GCTGTTCACCTCAAGATGGGCTGACCGCTGACGGGCGACAGTGTGGCTATGGCAT 1320
OY 1380 GAGCAGTACGATGAGCTGGAGGAGGCGCTGGGCGCACGTTGAGAGGAGTCTCGGCGCAT 1439
Db 1321 GAGCAGTACGATGAGCTGGAGGAGGCGCTGGGCGCACGTTGAGAGGAGTCTCGGCGCAT 1380
OY 1440 CGCCGCGCCCAACCGCTGGCTTCTGGCGCAGCTGCAGATCTACAGGGGCACTTGACG -- 1497
Db 1381 CGCCGCGCCCAACCGCTGGCTTCTGGCGCAGCTGCAGATCTACAGGGGCACTTGACGCGC 1440
OY 1498 ----- 1497
Db 1441 CAGCGCGCAGAGCAGATGCTGGAGCAGAAAGTGGGTGCTCCCGAGAGGACACC 1500
OY 1498 ----- 1521
Db 1501 AGCCCTGAGATCTTACACCATTTCCACCTTTCGCGCAGAACTGAGGGTGTGGGA 1560
OY 1522 GAGAGAGGTTGTAGGCTATGAGAGAGAGCAGGACGCGCGAAGAGAGAGCTGGGCGCAG 1581
Db 1561 GAGAGAGGTTGTAGGCTATGAGAGAGAGCAGGACGCGCGAAGAGAGAGCTGGGCGCAG 1620
OY 1582 GCCAGTATTAACCTCCGAGGGGCTCATGAGTTCATCACTTCTTGGAGCCTCTTGA 1641
Db 1621 GCCAGTATTAACCTCCGAGGGGCTCATGAGTTCATCACTTCTTGGAGCCTCTTGA 1680
OY 1642 GCTGAGAGCAGCTCAGAGAGCAGTATGTCAGAGGCTTCTTCCAGAGCTTC 1701
Db 1681 GCTGAGAGCAGCTCAGAGAGCAGTATGTCAGAGGCTTCTTCCAGAGCTTC 1740
OY 1702 ACATGAAGACCTTGCAGACCTTCCACAGCTTGCAGAGCAGGAGGAGGAGGAGGCT 1761
Db 1741 ACATGAAGACCTTGCAGACCTTCCACAGCTTGCAGAGCAGGAGGAGGAGGAGGAGGCT 1800
OY 1762 GAGCAGGGGCTTCAAGCTTGCCTGAAGTCCGCCAGTATGATGTTACCTCCAGGGCAG 1821
Db 1801 GAGCAGGGGCTTCAAGCTTGCCTGAAGTCCGCCAGTATGATGTTACCTCCAGGGCAG 1860
OY 1822 TGCCTGTGGTGCAGACCGGACCGGCTTCCAGAGGAGGAGGAGGAGGAGGAGGAGG 1881
Db 1861 TGCCTGTGGTGCAGACCGGACCGGCTTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1920
OY 1882 GCAGGAGAGGCTTGCATTTCTCTTACGCGCAGGTTCCGGAAGGTGTGAGACAGGCGAG 1941
Db 1921 GCAGGAGAGGCTTGCATTTCTCTTACGCGCAGGTTCCGGAAGGTGTGAGACAGGCGAG 1980
OY 1942 CGTGCATGACATGAGAGGAGGAGGCGAGGCTGAGGCTTACACATGCCAGCTCCCT 2001
Db 1981 CGTGCATGACATGAGAGGAGGAGGCGAGGCTGAGGCTTACACATGCCAGCTCCCT 2040
OY 2002 GACACTGAAGAGGATCCAACTCTGTGAGAAACACCTCAAGTGTGTGGCGGACACA 2061
Db 2041 GACACTGAAGAGGATCCAACTCTGTGAGAAACACCTCAAGTGTGTGGCGGACACA 2100
OY 2062 TTCTCTCAAGTCCGCGCCATACCGTCACTACAGCTTACCTCCACCCCTGTCTACTAC 2121
Db 2101 TTCTCTCAAGTCCGCGCCATACCGTCACTACAGCTTACCTCCACCCCTGTCTACTAC 2160
OY 2122 GGCCTACCTCCACCCCTGTCTACTACAGCTTACCTCTTACAGCTTTAAGTCCAGGCGC 2181
Db 2161 GGCCTACCTCCACCCCTGTCTACTACAGCTTACCTCTTACAGCTTTAAGTCCAGGCGC 2220
|||||

Oy	667	TTCAAGCCCATCTTCATCCAGACCATTGTGGCCACACTCCAGTATTGGACCAACATGT	726
Db	661	TTCAAGGCCCATCTTCACATCCAGACCATGT-----	688
Oy	727	GAGCAGCTGTAGCAGACGGCCTTTGATACGGGGTGGCAGTGGCCCTCACCGTGGGCCAGCCAC	786
Db	689	-----	688
Oy	787	TACCAAGAGAGACTGAACCTCCGAACAGAGCTGCCTCAATGAGTGGAGCGCTATGTGCGCAGC	846
Db	689	-----	688
Oy	847	CTGAGATCTCTGCGGCTCCCAACGCCGAGACCTGGCGGCTCTCAAGAACAGACATGT	906
Db	689	-----GCTCCTCAGAACAGAGCCAGTGT	711
Oy	907	GAGCAGGCGATCCGTCGTGTGAGCTGTGGAAAGTTGGATGTAGAGACCTGGAGAGTGTG	966
Db	712	GAGCAGGCGATCCGTCGTGTGAGCTGTGGAAAGTTGGATGTAGAGACCTGGAGAGTGTG	771
Oy	967	ACTTCCAAAGAGATCCGCGACAGGCTGTGGAGCTGCGGCTGTGGGCTCCCGCTCAGACATGC	1026
Db	772	ACTTCCAAAGAGATCCGCGACAGGCTGTGGAGCTGCGGCTCCCGCTCAGACATGC	831
Oy	1027	CGTACTTTCATCCAGCAACACAGATGCTGCTGTGTGTGACACCGGAGACCGACTCCCGC	1086
Db	832	CGTACTTTCATCCAGCAACACAGATGCTGCTGTGTGTGACACCGGAGCCGAGCTCCCGC	891
Oy	1087	ATTTCTCCCGCCACCTCTACTCTGGGCTCAGAGTGAAGCGCAGCAACCTGAGAGAGCTGCAG	1146
Db	892	ATTTCTCCCGCCACCTCTACTCTGGGCTCAGAGTGAAGCGCAGCAACCTGAGAGAGCTGCAG	951
Oy	1147	AGGAACAGGGTCACCCACATCTTGAACATGTGCCCGGAGATTGACACTTTCACCTGAG	1206
Db	952	AGGAACAGGGTCACCCACATCTTGAACATGTGCCCGGAGATTGACACTTTCACCTGAG	1011
Oy	1207	CGCTTCACATTCACACAATGTGCGCTCTGTGGATGAGAGATCGGCCAGCTGCTCGGCAC	1266
Db	1012	CGCTTCACATTCACACAATGTGCGCTCTGTGGATGAGAGATCGGCCACCTGCTCGGCAC	1071
Oy	1267	TGGAAAGAGAGCGACCGCTTCATTTGAGGGTCGCAAGAGCACAGGGACCCACAGTGTGCTG	1326
Db	1072	TGGAAAGAGAGCGACCGCTTCATTTGAGGGTCGCAAGAGCACAGGGACCCACAGTGTGCTG	1131
Oy	1327	CACGTCAAGATGGGCGCTCAGCCGCTCAGCGGCCACAGTGTGCTGCTATGCCATGAAGCAG	1386
Db	1132	CACGTCAAGATGGGCGCTCAGCCGCTCAGCGGCCACAGTGTGCTGCTATGCCATGAAGCAG	1191
Oy	1387	TACCAATGCACTCGAGAGAGGCGCTGCGCCACAGTGTGAGAGACTCCGCGCCATCGCCCGC	1446
Db	1192	TACCAATGCACTCGAGAGAGGCGCTGCGCCACAGTGTGAGAGACTCCGCGCCATCGCCCGC	1251
Oy	1447	CCCAACCTGGCTTCCTGCGCCACAGCTGCAGATCTACAGGGGATCCTGACGGCCAGAAC	1506
Db	1252	CCCAACCTGGCTTCCTGCGCCACAGCTGCAGATCTACAGGGGATCCTGACGGCCAGAAC	1311
Oy	1507	TGAGGCTGTGGGAGAGAGGTTGTAGGCATGGAAGAGACCAGCGACGCCCGAAGA	1566
Db	1312	TGAGGCTGTGGGAGAGAGGTTGTAGGCATGGAAGAGACCAGCGACGCCCGAAGA	1371
Oy	1567	AGAGCCCTGGG--CCAGCGCCACGTATAAACCTCCAGGGGCTATAGATTCATCAGTCTT	1624
Db	1372	AGAGCCCTGGG-----CCAGCGCCACGTATAAACCTCCAGGGGCTATAGATTCATCAGTCTT	1431
Oy	1625	CTGAGAGCCCTCTT--GGAGCTGTGAGAGACACTCAG--AGACAGTGTACATGGCAGAGCTCT	1682
Db	1432	CTGAGAGCCCTCTTCTTGGAGGCTGTGAGAGACACTCAGTATAGACAGTATGCCAAGAGTCT	1491
Oy	1683	TCTCTTCCACAGAGTCTTCATCATGAAGAGCTCTGACGCCCTTCCACAGCTTGCAAGA	1742
Db	1492	TCTCTTCCACAGAGTCTTCATCATGAAGAGCTCTGACGCCCTTCCACAGCTTGCAAGA	1551

QY	1743	CCAGAGGAGGCCACAGTGTGGACAGGGGGGCTCTAGCCTCGCTGAAGTCCGCCACGTAG	1802
Db	1552	CCAAAGGAGGGCCACAGTGTGGACAGGGGGGCTCTAGCCTCGCTGAAGTCCGCCACGTAG	1611
QY	1803	TGTGTACCTCTCAAGGGACAGTGTGGTGGCCAAACGGAGCCAGAGGCTCCAGAGACAGG	1862
Db	1612	TGTGTACCTCTCAAGGGACAGTGTGGTGGCCAAACGGAGCCAGAGGCTCCAGAGACAGG	1671
QY	1863	AGCAAGGGCAGAGGGCAGAGGGCAGAGAGAGCCCTGCATTTCTCTAAGCCCAAGTTCGCGA	1922
Db	1672	AGCAAGGGCAGAGGGCAGAGGGCAGAGAGAGCCCTGCATTTCTCTAAGCCCAAGTTCGCGA	1731
QY	1923	AGGTGTGTAGACAGGCCACGCGTGCATGCAGTGTGAAGAGAGAGGGCGAGCGCTGACCTCA	1982
Db	1732	AGGTGTGTGTAGACAGGCCACGCGTGCATGCAGTGTGAAGAGAGAGGGCGAGCGCTGACCTCA	1791
QY	1983	CACATGGCCACAGCTCCGCTCTACACATGAAGAGAGTCCACAACTCTGTGGAGAAACACCTC	2042
Db	1792	CACATGGCCACAGCTCCGCTCTACACATGAAGAGAGTCCACAACTCTGTGGAGAAACACCTC	1851
QY	2043	ACGTCGTGTGGCGACACATTTCCCTCAGTCCCGCCCACTACCCGTCACATACAGCTCAC	2102
Db	1852	ACGTCGTGTGGCGACACATTTCCCTCAGTCCCGCCCACTACCCGTCACATACAGCTCAC	1911
QY	2103	CTCCACACCCCTGTACATGAGGGCTCACCTCCACACCCCTGTACATACAGCTCACCTCTA	2162
Db	1912	CTCCACACCCCTGTACATGAGGGCTCACCTCCACACCCCTGTACATACAGCTCACCTCTA	1971
QY	2163	CAGCCTTAAGTCCCAAGGGCCATGTCTGCTCTGTCCAAGGGCTCAAGACTTTCTAAGTGGGA	2222
Db	1972	CAGCCTTAAGTCCCAAGGGCCATGTCTGCTCTGTCCAAGGGCTCAAGACTTTCTAAGTGGGA	2031
QY	2223	TGTGTGTAGAGGACTGAAGGTACCTTTGGGGGGACAAGAGACCTTAGTTCATTTCTCACT	2282
Db	2032	TGTGTGTAGAGGACTGAAGGTACCTTTGGGGGGACAAGAGACCTTAGTTCATTTCTCACT	2091
QY	2283	CTAGCCCTGCACACTGCCTGTGGACAGAGTAAGAAACAGAGCTTCCCGTGCAGAAAGGG	2342
Db	2092	CTAGCCCTGCACACTGCCTGTGGACAGAGTAAGAAACAGAGCTTCCCGTGCAGAAAGGG	2151
QY	2343	TCAGGCTCCCAACCCCGCCGCTCCCTGTGACTCCTGTCTCTCTCCAGTTATTTCTGTGG	2402
Db	2152	TCAGGCTCCCAACCCCGCCGCTCCCTGTGACTCCTGTCTCTCTCCAGTTATTTCTGTGG	2211
QY	2403	AACCAAGCCAGGACAGGACACAGTGGGGCCCAAGGACAGGAGATCCACAGGCCCAAGC	2462
Db	2212	AACCAAGCCAGGACAGGACACAGTGGGGCCCAAGGACAGGAGATCCACAGGCCCAAGC	2271
QY	2463	CGCGGGAGGCTGGAAGGGCTGGGAGATCGCTTCCCTCATCTCACCTCCACACCGGTCAGGTC	2522
Db	2272	CGCGGGAGGCTGGAAGGGCTGGGAGATCGCTTCCCTCATCTCACCTCCACACCGGTCAGGTC	2331
QY	2523	TTTGTGCTGTGCCCAACACTCCTGTGTAGACACACGCGCAATATACAGGGCACAGGCCAGA	2582
Db	2332	TTTGTGCTGTGCCCAACACTCCTGTGTAGACACACGCGCAATATACAGGGCACAGGCCAGA	2391
QY	2583	GATAGTCTCTTTTGTGTCTTTCTGGGCGCTGTGCTAGTGTAGTTTTTCATAGGCTTACAGT	2642
Db	2392	GATAGTCTCTTTTGTGTCTTTCTGGGCGCTGTGCTAGTGTAGTTTTTCATAGGCTTACAGT	2451
QY	2643	ATCTGGCTTTGTACTGTGAAGAAATAAACACATTTTCATATTGTGGTTAAAAA	2702
Db	2452	ATCTGGCTTTGTACTGTGAAGAAATAAACACATTTTCATATTGTGGTTAAAAA	2511
QY	2703	AA 2704	
Db	2512	AA 2513	
RESULT 8			
ABL40805			
ID ABL40805 standard; cDNA; 2322 BP.			
XX			

AC ABL40805;
XX
DT 03-JUL-2002 (first entry)
XX
XX Human MAP kinase phosphatase-like enzyme encoding cDNA.
DE
XX Mitogen activated protein; MAP kinase phosphatase-like enzyme;
XX antidiabetic; antidiabetic; anorectic; cytosolic; cardiac; human;
KW antiparkinsonian; cerebroprotective; neuroprotective; nootropic; gene;
KW neoplastic; anticonvulsant; anti-HIV; antiarrhythmic; hypotensive;
KW antiallergic; dermatological; vulnery; gene therapy; ss.
XX
OS Homo sapiens.
XX
FH Location/Qualifiers
FT CDS
FT 1..2322
FT /tag= a
FT /product= "MAP kinase phosphatase-like enzyme"
FT /note= "contains internal codon deletions"
FT /transl_except= "(pos: 156..157, aa: Leu)"
FT /note= "there is an apparent one codon deletion which
FT alters the reading frame"
FT /transl_except= "(pos: 180..181, aa: Arg)"
FT /note= "there is an apparent one codon deletion which
FT alters the reading frame"
FT /transl_except= "(pos: 414..415, aa: Val)"
FT /note= "there is an apparent one codon deletion which
FT alters the reading frame"
FT /transl_except= "(pos: 1197..1198, aa: Tyr)"
FT /note= "there is an apparent one codon deletion which
FT alters the reading frame"
FT /transl_except= "(pos: 2088..2089, aa: Xaa)"
FT /note= "there is an apparent one codon deletion which
FT alters the reading frame"
XX
PN MO200220732-A2.
XX
PD 14-MAR-2002.
XX
PF 27-AUG-2001; 2001MO-EP09848.
XX
PR 07-SEP-2000; 2000US-230709P.
XX
PA (FARB) BAYER AG.
XX
PI Liou J;
XX
XX WPI: 2002-339802/37.
DR P-PSDB; ABB07845.
XX
XX New human mitogen activated protein kinase phosphatase-like enzyme
PT polypeptide, regulators of which are useful for preventing treating
PT allergies including asthma, diabetes, obesity, cancer and
PT cardiovascular diseases
XX
XX Claim 1; Fig 13; 134pp: English.
XX
XX The invention relates to a purified human mitogen activated protein (MAP)
CC kinase phosphatase-like enzyme polypeptide. The enzyme can be expressed
CC by standard recombinant methodology. The MAP kinase phosphatase-like
CC enzyme and encoding polynucleotides are useful for screening for
CC modulators which are used for treating a MAP kinase phosphatase-like
CC enzyme dysfunction related disease, such as asthma, a central nervous
CC system disorder, diabetes, obesity, chronic obstructive pulmonary
CC disease, cancer or a cardiovascular disease. The enzyme can be regulated
CC to treat allergies including asthma, allergic rhinitis, atopic
CC dermatitis, and anaphylaxis, central nervous system disorders such as
CC brain injuries, Parkinson's disease, dementia, multiple sclerosis,
CC stroke, Alzheimer's disease, Huntington's disease, schizophrenia, Pick's
CC disease, Creutzfeldt-Jacob dementia, progressive nuclear palsy, and human
CC immunodeficiency virus (HIV) dementia, and cardiovascular diseases
CC including myocardial infarction, ischemic diseases of the heart, atrial
CC and ventricular arrhythmia, hypertensive vascular diseases and peripheral

CC vascular diseases. The enzyme is useful in diagnostic assays for
CC detecting diseases and abnormalities or susceptibilities to diseases or
CC abnormalities related to the presence of mutations in the encoding
CC nucleic acid sequences. The present sequence represents the human MAP
CC kinase phosphatase-like enzyme polypeptide encoding cDNA.
XX
S0 Sequence 2322 BP: 467 A; 751 C; 718 G; 386 T; 0 other;
Query Match 62.8%; Score 1699; DB 24; Length 2322;
Best Local Similarity 90.9%; Pred. No. 0;
Matches 1987; Conservative 0; Mismatches 75; Indels 123; Gaps 12;
QY 132 CGGCGCCCTCCACGCGCCGCGGCGCTGGGACCGAGCCAGGCGAAGAGCTGCTCA 191
DB 147 GCGCCCTCACCGCTGGGCTCTCTCTGCGCAGAGACTGTGTCAGAGAGAGTCACTCA 206
QY 192 GCGAAGCAGACGCTTGGCGTGTCTCC-GTGGGCTGTCTCTGGAGTCCAGATGAGAGG 250
DB 207 GCGA----AGAGCTTGGCGTGTCTCTCTGGGCTGTCTCTGGAGTCCAGATGAGAGG 262
QY 251 ACAATGATGATGCACAGAGGCGCACTTGTGAGCCACAGAGAGAGGCCCGGAGTGAAGG 310
DB 263 ACAATGATGATGCACAGAGGCGCACTTGTGAGCCACAGAGAGAGGCCCGGAGTGAAGG 322
QY 311 AGCTCCAGCGGG---ACGAGACAGACTT--GGGGCAAGGATCCAGAGTCCGAGAGCA 365
DB 323 AGCTCCAGCGGGGAAACCGACAGACTTGTGTCGTCAGAGATCCAGATCCCGAGAGCA 382
QY 366 GGAGAGCAGAGGAGCAGACCTGACCTCATGCTGACAGCTGCTGAGAGCGGCGAGATGAT 425
DB 383 GGAGAGCAGAGGAGCAGACCTGACCTCATGAGCTGCTGAGAGCGGCGAGATGATGAT 442
QY 426 CCGCCTGGCA--GCCAGCTGAGAGGACCCCGGCTCCCGGCTCCGCTACCTGC----- 478
DB 443 CCGACTTGGAAAGCCCAAGCTGAGAGGACCCCGGCTCCCGGATCCGATACCTTGT 502
QY 479 TGGTAGTTCTACAGAGAGAG--GAGAGGCTGAG--CCAGATGAGAGGCTCTCTGAGG 536
DB 503 TGGTAGTTCTACAGAGAGAGAGAGAGAGTGTGAGCCAGATGAGAGAGTCTCTGAGG 562
QY 537 -COTGATTTCCCTGACAGCAGCTCCCGCAGCTGCACCCCTGGGCTGTCTTGGCCCTCT 595
DB 563 AGGTGATTTCCCTGACAGCAGCTCCCGCAGCTGCACCCCTGGGCTGTCTTGGCCCTCT 622
QY 596 GGAGTACACCCAGAGTGTACTAGATGAGAGCGGGGCTTCAAGCGTGTGGGAGGCG 655
DB 623 GGAGTACACCCAGAGTGTACTAGATGAGAGCGGGGCTTCAAGCGTGTGGGAGGCG 682
QY 656 AAGCGGATCTTCAAGCCCATCTCCATCCAGACATGTGGGCGACACTCCAGGTATTGC 715
DB 683 AAGCGGATCTTCAAGCCCATCTCCATCCAGACATGTGGGCGACACTCCAGGTATTGC 742
QY 716 ACCAGCATGTGAGGAGGCTTAGAGGCGGCTTGTACCGGGTGGAGTCCCTCACT 775
DB 743 ACCAGCATGTGAGGAGGCTTAGAGGCGGCTTGTACCGGGTGGAGTCCCTCACT 802
QY 776 GGGCCAGCCACTACAGAGAGAGACTGAAGTCCGAAAGAGTGTCCATGTGAGTGAAGG 835
DB 803 GGGCCAGCCACTACAGAGAGAGACTGAAGTCCGAAAGAGTGTCCATGTGAGTGAAGG 862
QY 836 CTATGGCGGCTGGAGTCTCTGCGGCTCCAGGCCAGAGCTTGCGGGTCTCTCAAG 895
DB 863 CTATGGCGGCTGGAGTCTCTGCGGCTCCAGGCCAGAGCTTGCGGGTCTCTCAAG 922
QY 896 AGGAGCATGAGAGAGGCGATCCGTGAGTGTGAAAAGTGTGATGTGATGATGATGAC 955
DB 923 AGGAGCATGAGAGAGGCGATCCGTGAGTGTGAAAAGTGTGATGTGATGATGATGAC 982
QY 956 TGGAGATGCTACTTCCAAAGATATCCAGAGCTGTGAGTGTGAGTGTGAGTGTGAGT 1015
DB 983 TGGAGATGCTACTTCCAAAGATATCCAGAGCTGTGAGTGTGAGTGTGAGTGTGAGT 1042
QY 1016 TCCAGCATGATCCGTAAGTCTATGACAAACAGATGTGCTGTGTGAGCAGAGCGGAGC 1075

Db	1043	TCGACGCTGACCTGACCTTCATGACAAACACAGTGTGCTGCTGTGGCACAAGCGGAGCC	1102
Qy	1076	GAGCCTCCGCGATCTTCCCCACACCTTACCTGAGGCTCAGAGTGAACACAGCAAACTGG	1135
Db	1103	GAGCCTCCGCGATCTTCCCCACACCTTACCTGAGGCTCAGAGTGAACAGCAAACTGG	1162
Qy	1136	AGGAGCTGCAAGAGGAACAGGGTACCCACATCTTGAAACATGGCCCGGAGATTGACAAC	1195
Db	1163	AGGAGCTGCAAGAGGAACAGGGTACCCACATCTTG--ATGGCCGGAGATTGACAAC	1219
Qy	1196	TCGACCTGAGCGCTTACCTTACCAAAATGTGGGCTTGGATGAGGATGGGCGGAC	1255
Db	1220	TCGACCTGAGCGCTTACCTTACCAAAATGTGGCTTGGATGAGGATGGGCGGAC	1279
Qy	1256	TGCTGCGCACTGGAGAGAGACGACCGCTTCAATTGAGGCTGCAAGACACAGGACCC	1315
Db	1280	TGCTGCGCACTGGAGAGAGACGACCGCTTCAATTGAGGCTGCAAGACACAGGACCC	1339
Qy	1316	ACGTCGTGTCACCTGCAAGATGGGGGTAGCGCGTCACGGCCGCAAGTCGTCATG	1375
Db	1340	ACGTCGTGTCACCTGCAAGATGGGGGTAGCGCGTCACGGCCGCAAGTCGTCATG	1399
Qy	1376	CCATGAAAGCAGTACGATGCAAGCTGAGCAGGACCCCTGGCCACGTGCAAGAGCTCGGC	1435
Db	1400	CCATGAAAGCAGTACGATGCAAGCTGAGCAGGACCCCTGGCCACGTGCAAGAGCTCGGC	1459
Qy	1436	CCATCGCGCGCCCAACCCCTGGCTTCTTCGCGCAGCTGCAATCTACCAAGGCATCTGA	1495
Db	1460	CCATCGCGCGCCCAACCCCTGGCTTCTTCGCGCAGCTGCAATCTACCAAGGCATCTGA	1519
Qy	1496	CG-----	1497
Db	1520	CGCGACACCGCCACAGACCATGTCGTGGGACAGAAAGTGGGTGTCTCCCAAGAGAC	1579
Qy	1498	-----GCCAGACCTGAGGCTGTG	1517
Db	1580	ACCCAGCCCTGGAAGTCTCTACACATTCCTCCACTCTTCCGCGCAACACTGGAGGTGATG	1639
Qy	1518	GGGAGGAGAGGTTGTAGGATGGAAGAGAGGACGACCCCGAAAGAGAGCTGGGC	1577
Db	1640	GGGAGGAGAGGTTGTAGGATGGAAGAGAGGACGACCCCGAAAGAGAGCTGGGC	1699
Qy	1578	CACGGCACGATTAACCTTCCGAGGGGTGATGAGTCCATCAGTCTTTCGAGCCCTCT	1637
Db	1700	CACGGCACGATTAACCTTCCGAGGGGTGATGAGTCCATCAGTCTTTCGAGCCCTCT	1759
Qy	1638	TGGAGCTGGGAGACACTCTCAGAACCCAGTGAATGTCAGAGTCTTCTTCCACAGT	1697
Db	1760	TGGAGCTGGGAGACACTCTCAGAACCCAGTGAATGTCAGAGTCTTCTTCCACAGT	1819
Qy	1698	CTTCACATGAAGAGCCTCTGCAAGCTTCCACAGCTTGCAAGGACCAAGGAGGACGAC	1757
Db	1820	CTTCACATGAAGAGCCTCTGCAAGCTTCCACAGCTTGCAAGGAGGACGAC	1879
Qy	1758	AGGTGACACAGGGGGCTCAGACCTGCCCCTAAGTCCCGCCAGTCAATGTTACCTCCAG	1817
Db	1880	AGGTGACACAGGGGGCTCAGACCTGCCCCTAAGTCCCGCCAGTCAATGTTACCTCCAG	1939
Qy	1818	GCAATGCGCTGTGTGCAACCGGACCCAGGACCTTCAGAGCAGGAGCAGGGGAC	1877
Db	1940	GCAATGCGCTGTGTGCAACCGGACCCAGGACCTTCAGAGCAGGAGCAGGGGAC	1999
Qy	1878	AGGGGACAGGAGACCCCTGCAATTCCTCTACGCCACAGTTCCGGAAGTGTGAGACAG	1937
Db	2000	AGGGGACAGGAGACCCCTGCAATTCCTCTACGCCACAGTTCCGGAAGTGTGAGACAG	2059
Qy	1938	CCACGAGTCAAGTACAGTGGAGAGAGGGGAGGCGCGAAGCCCTCACATGCGCACGCTC	1997
Db	2060	CCACGAGTCAAGTACAGTGGAGAGAGGGG--GGCTGAGACCTCTCACATGCGCACGCTC	2117
Qy	1998	CCCTGCACTGAAGAGATTCACAACTCCTTGGAGAAACACCTCAACCTGTGTGCCCA	2057

Db	2118	CCGAGACACGTAAAGATATCCACAACTCCTTGGAGAAACACCTCAGCGTCTGTGGCCGA	2177		
Qy	2058	CACATTCCTCTCAGCTCCGGCCCATPACCCGTCACTACAGCTTACCTCCACCCCTGTCA	2117		
Qy	2118	CTAGGGCTTACCTCCACCCCGTGTACTACTACAGCGTACCTCTCTAAGCCTTAAGTCCCA	2177		
Qy	2178	GGCCCATGTCTGCTGTCCAAAGGC	2202		
RESULT 9					
AAH14722					
ID	AAH14722	standard; cDNA; 1755 BP.			
XX					
AC	AAH14722;				
XX					
DT	26-JUN-2001	(first entry)			
XX					
DE	Human cDNA sequence SEQ ID NO:12452.				
XX					
KW	Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss				
XX					
OS	Homo sapiens.				
XX					
PN	EP1074617-A2.				
XX					
PD	07-FEB-2001.				
XX					
PE	28-JUL-2000; 2000EP-0116126.				
XX					
PR	29-JUL-1999; 99JP-0248036.				
XX					
PR	27-AUG-1999; 99JP-0300253.				
XX					
PR	11-JAN-2000; 2000JP-0118776.				
XX					
PR	02-MAY-2000; 2000JP-0183767.				
XX					
PR	09-JUN-2000; 2000JP-0241899.				
XX					
PA	(HELI-) HELIX RES INST.				
XX					
PI	Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;				
PI	Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;				
XX					
DR	WPI; 2001-318749/34.				
XX					
PT	Primer sets for synthesizing polynucleotides, particularly the 5602				
XX					
PT	full-length cDNAs defined in the specification, and for the detection				
XX					
PT	full-length cDNAs -				
XX					
PS	Claim 8; SEQ ID 12452; 2537bp + CD ROM; English.				
XX					
CC	The present invention describes primer sets for synthesizing 5602				
CC	full-length cDNAs defined in the specification. Where a primer set				
CC	comprises: (a) an oligo-dT primer and an oligonucleotide complementary				
CC	to the complementary strand of a polynucleotide which comprises one of				
CC	the 5602 nucleotide sequences defined in the specification, where the				
CC	oligonucleotide comprises at least 15 nucleotides; or (b) a combination				
CC	of an oligonucleotide comprising a sequence complementary to the				
CC	complementary strand of a polynucleotide which comprises a 5'-end				
CC	sequence and an oligonucleotide comprising a sequence complementary to a				
CC	polynucleotide which comprises a 3'-end sequence, where the				
CC	oligonucleotide comprises at least 15 nucleotides and the combination of				
CC	the 5'-end sequence/3'-end sequence is selected from those defined in				
CC	the specification. The primer sets can be used in antisense therapy and				
CC	in gene therapy. The primers are useful for synthesizing polynucleotides				
CC	particularly full-length cDNAs. The primers are also useful for the				
CC	detection and/or diagnosis of the abnormality of the proteins encoded by				
CC	the full-length cDNAs. The primers allow obtaining of the full-length				
CC	cDNAs easily without any specialised methods. AAH03156 to AAH13628 and				

CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
XX of the present invention.

Sequence 1755 BP; 350 A; 577 C; 505 G; 323 T; 0 other:

Query Match 49.4%; Score 1337; DB 22; Length 1755;
Best Local Similarity 94.7%; Pred. No. 1.9e-258;
Matches 1427; Conservative 0; Mismatches 0; Indels 80; Gaps 1;

```

OY 1174 ATGGCCCGGAGATTGACAACTTACCTTGAGCGCTTCACTTCAACAATGTGGCCTC 1233
DB 322 ATGGCCCGGAGATTGACAACTTACCTTGAGCGCTTCACTTCAACAATGTGGCCTC 381
OY 1234 TGGGATGAGAGTGGGGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1293
DB 382 TGGGATGAGAGTGGGGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 441
OY 1294 GCTGCAAGACACAGGACACCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1353
DB 442 GCTGCAAGACACAGGACACCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 501
OY 1354 GCGGCCACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1413
DB 502 GCGGCCACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 561
OY 1414 GCGCCAGTGGAGAGTCCGCGCCATGCGCGCCCAACCTGGCTTCCGCGCCAGCTG 1473
DB 562 GCGCCAGTGGAGAGTCCGCGCCATGCGCGCCCAACCTGGCTTCCGCGCCAGCTG 621
OY 1474 CAGATCTACAGGAGCATCTGACGGCCAGAACTGAGGGTGTGGGAGAGAAAGTTGT 1533
DB 622 CAGATCTACAGGAGCATCTGACGGCCAGAACTGAGGGTGTGGGAGAGAAAGTTGT 681
OY 1534 AGGATGGAAGAGAGCCAGGACGCGCGGAAAGAGCTGGGCGACGGCCAGCTATAAA 1593
DB 682 AGGATGGAAGAGAGCCAGGACGCGCGGAAAGAGCTGGGCGACGGCCAGCTATAAA 741
OY 1594 CTTCCGAGGGGTGATGAGTGCATGATCTTCTGAGAGCCCTTCTTGGAGCTGGAGAC 1653
DB 742 CTTCCGAGGGGTGATGAGTGCATGATCTTCTGAGAGCCCTTCTTGGAGCTGGAGAC 801
OY 1654 CTGAGAGACAGTACATGTCAGAGGTCTTCTCTCCACAGAGTCTTACATGAGAGCC 1713
DB 802 CTGAGAGACAGTACATGTCAGAGGTCTTCTCTCCACAGAGTCTTACATGAGAGCC 861
OY 1714 TCTGCAAGCCTTCCACAGCTTTCAGAGGACCAAGGAGGCGCAGCAGGAGGCC 1773
DB 862 TCTGCAAGCCTTCCACAGCTTTCAGAGGACCAAGGAGGCGCAGCAGGAGGCC 921
OY 1774 TCAGCCCTGAGTGAAGTCCCGCCAGTACAGTGTACCTTCCAGGGCAGTCCCGTGGG 1833
DB 922 TCAGCCCTGAGTGAAGTCCCGCCAGTACAGTGTACCTTCCAGGGCAGTCCCGTGGG 981
OY 1834 CAACCGGACCCAGGCTTTCAGAGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGCC 1893
DB 982 CAACCGGACCCAGGCTTTCAGAGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGCC 1041
OY 1894 CTGCAATTTCTTACGCGCCAGGTTCCGGAAGGTGTGAGACAGGCGCAGCTGCATG 1953
DB 1042 CTGCAATTTCTTACGCGCCAGGTTCCGGAAGGTGTGAGACAGGCGCAGCTGCATG 1101
OY 1954 TGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2013
DB 1102 TGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1161
OY 2014 GATCCACAGTCTTGGAGAAACACCTTCACTGTCTGTTGGCGACACATTTCTCTACGT 2073
DB 1162 GATCCACAGTCTTGGAGAAACACCTTCACTGTCTGTTGGCGACACATTTCTCTACGT 1221
OY 2074 CCGCCCATACCGTCTACTACAGCTTCACTGCCACCTGTGTACATGAGGCTTACCTCC 2133
DB 1221 CCGCCCATACCGTCTACTACAGCTTCACTGCCACCTGTGTACATGAGGCTTACCTCC 2181
```

```

DB 1222 CCGCCCATACCGTCTACTACAGCTTCACTGCCACCTGTGTACATGAGGCTTACCTCC 1281
OY 2134 CACCCCTGTCACTACAGCTTCACTACCTTCACTACCTTAACTCCAGGCCATGTGCTGCTG 2193
DB 1282 CACCCCTGTCACTACAGCTTCACTACCTTCACTACCTTAACTCCAGGCCATGTGCTGCTG 1341
OY 2194 TCCAGGGCTCAAGACTTCTTAACTGGGATGTGTAGAGGAGACTAAGGTACTTTGGGG 2253
DB 1342 TCCAGGGCTCAAGACTTCTTAACTGGGATGTGTAGAGGAGACTAAGGTACTTTGGGG 1401
OY 2254 GCACAGCAGCCTTACTTGTATTTCTACCTTACCTTACCTTACCTTACCTTACCTTAC 2313
DB 1402 GCACAGCAGCCTTACTTGTATTTCTACCTTACCTTACCTTACCTTACCTTACCTTAC 1434
OY 2314 TGAACACAGAGTCCCGTGCAGAAAAAGGGTCAAGGCTCCACCCCGCCCTTCCCTGCA 2373
DB 1435 -----CCCTGGA 1441
OY 2374 CCTCTGTCTCTTCCAGTTTCACTTCTGAGAACCCAGGACGAGGACCAAGTGGCCCTC 2433
DB 1442 CCTCTGTCTCTTCCAGTTTCACTTCTGAGAACCCAGGACGAGGACCAAGTGGCCCTC 1501
OY 2434 AAGGAGGAGGAGATCTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2493
DB 1502 AAGGAGGAGGAGATCTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1561
OY 2494 TCCCTCATCAGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGG 2553
DB 1562 TCCCTCATCAGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGG 1621
OY 2554 CACGCGAGATCACAGGAGGACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2613
DB 1622 CACGCGAGATCACAGGAGGAGGACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1681
OY 2614 GGTAGTCACTTTTTCATAGCTTACAGTATCTGCTTGTGCTGAGAAATAAACACAT 2673
DB 1682 GGTAGTCACTTTTTCATAGCTTACAGTATCTGCTTGTGCTGAGAAATAAACACAT 1741
OY 2674 TTTTCATA 2680
DB 1742 TTTTCATA 1748

RESULT 10
ID ABL40801
ABLA0801 standard; DNA: 1755 BP.
XX
AC ABL40801:
XX
XX 03-JUL-2002 (first entry)
XX
DE Human MAP kinase phosphatase-like enzyme DNA fragment.
XX
XX Mitogen activated protein; MAP; MAP kinase phosphatase-like enzyme;
XX antidiabetic; antidiabetic; anorectic; cytoprotective; human;
XX antiparkinsonian; cerebroprotective; neuroprotective; nootropic; gene;
XX neuroleptic; anticonvulsant; anti-HIV; antiarrhythmic; hypotensive;
XX antiallergic; dermatological; vulnery; gene therapy; ds.
XX
OS Homo sapiens.
XX
XX PN WO200220732-A2.
XX
XX PD 14-MAR-2002.
XX
XX PF 27-AUG-2001; 2001WO-EP09848.
XX
XX PR 07-SEP-2000; 2000US-230709P.
XX
XX PA (FARB ) BAYER AG.
XX
XX Llou J;
XX
```


AC ABNS9832;
XX
DT 28-JUN-2002 (first entry)
XX
DE Novel human coding sequence SEQ ID NO: 243.
XX
XX Human; antinaemic; vulnery; antinflammatory; immunomodulator;
KW antifertility; cerebroprotective; cytoskeletal; rheumatic; gene therapy;
KW neuroprotective; antiparkinsonian; protein therapy; EST;
KW expressed sequence tag; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200222660-A2.
XX
XX 21-MAR-2002.
PD 10-SEP-2001; 2001WO-US26015.
PF 11-SEP-2000; 2000US-0659671.
PR (HYSE-) HYSEQ INC.
XX
PA Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX
XX MPI: 2002-292408/33.
DR P-PSDB: ABB97419.
XX
XX An isolated polynucleotide for treating diseases associated with its
PT encoded polypeptide such as cancer and multiple sclerosis -
XX
XX Claim 1; SEQ ID NO 243; 509bp; English.
XX
CC The present invention provides the protein and coding sequences of 444
CC novel human proteins. These were isolated from expressed sequences tags
CC (ESTs). They can be used to stimulate cell growth, to regulate
CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
CC e.g. in burn treatment, to regulate the immune system e.g. to treat
CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
CC infertility, to regulate haemostasis or thrombolysis e.g. to treat
CC stroke and cancer, to screen for drugs, to treat inflammatory conditions
CC e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.
CC Parkinson's disease. The present sequence is a coding sequence of the
CC invention.
XX
XX
SQ Sequence 2061 BP; 415 A; 672 C; 605 G; 369 T; 0 other;
Query Match 47.4%; Score 1280.8; DB 24; Length 2061;
Best Local Similarity 73.8%; Pred. No. 3.6e-247;
Matches 1995; Conservative 0; Mismatches 12; Indels 697; Gaps 1;
1 CGCTCTCTGCTGCTGCGGGTCCAGACTGTCGCGGGGTTAGGAGGAGCGCTGCC 60
DB 54 CGTCTTCTCTGCTGCGGGTCCAGACTGTCGCGGGGTTAGGAGGAGCGCTGCC 113
QY 61 CGGTCCAGCCAGGTGCTCGCGGCTGCTCCATGGCCCTGTCAAGTGAAGCCGTTG 120
DB 114 CGGTCCAGCCAGGTGCTCGCGGCTGCTCCATGGCCCTGTCAAGTGAAGCCGTTG 173
QY 121 CCCCCGGGACGCGGCGCTTCACGCGCGCTGCGGAGCCGAGGCGGCGCAAG 180
DB 174 CCCCCGGGACGCGGCGCTTCACGCGCGCTGCGGAGCCGAGGCGGCGCAAG 233
QY 181 AGTCGACTCCAGGAGGAGGAGCTTGGCGGTCTCCGTTGGGCTGCTCCGAGTGCAG 240
DB 234 AGTCGACTCCAGGAGGAGGAGCTTGGCGGTCTCCGTTGGGCTGCTCCGAGTGCAG 293
QY 241 GATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300
DB 294 GATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 353
QY 301 AGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360

DB 354 AGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 413
QY 361 AAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420
DB 414 AAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 473
QY 421 GACATCCGCTGAGAGCCAGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480
DB 474 GACATCCGCTGAGAGCCAGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 533
QY 481 GTAGTTTCTACACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540
DB 534 GTAGTTTCTACACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 593
QY 541 GATTTCCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600
DB 594 GATTTCCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 653
QY 601 GACACCCAGGCTGACTTATGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660
DB 654 GACACCCAGGCTGACTTATGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 713
QY 661 CGGATCTTCAAGCCCATCTCCATCCAGACATGTGGGCAACATCTGAGGAGGAGGAGG 720
DB 714 CGGATCTTCAAGCCCATCTCCATCCAGACATGTGGGCAACATCTGAGGAGGAGGAGG 773
QY 721 GCATGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
DB 774 GCATGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 833
QY 781 AGCCACTACAGAGAGAGAGTCACTCCAGACAGAGGAGGAGGAGGAGGAGGAGGAGG 840
DB 834 AGCCACTACAGAGAGAGAGTCACTCCAGACAGAGGAGGAGGAGGAGGAGGAGGAGG 893
QY 841 GCGGACCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 900
DB 894 GCGGACCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 953
QY 901 CGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960
DB 954 CGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 992
QY 961 AGTGTACTTCCAAAGAGATCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1020
DB 993 ----- 992
QY 1021 CAGTAACCTGACTTATGAGAGACAGAGATGCTGCTGTTGGAGCAGAGGAGGAGGAGG 1080
DB 993 ----- 992
QY 1081 TCCCGCATCTTCCCGCACCCTTACCTGAGGCTCAGAGTGAAGGAGCAAGACCTGAGAG 1140
DB 993 ----- 992
QY 1141 CTGACAGAGGAGAGGAGTCAACCATCTTTGAACATGGCCCGGAGATGACAACTTCTAC 1200
DB 993 ----- 992
QY 1201 CCTGAGCGCTTACCTACCAACATATGCGGCTCTGAGATGAGAGTGGCCAGCTGCTG 1260
DB 993 ----- 992
QY 1261 CCGCACTGGAAGAGAGAGCAGCCGCTTCAATTGAGGCTGCAAGAGCAGAGGAGCAGCAG 1320
DB 993 ----- 992
QY 1321 CTGTGCTCACTGCAAGATGGGCTCAGCGGCTAGCGGCCACAGTCTCTGCTATGCTATG 1380
DB 993 ----- 992
QY 1381 AAGCACTACGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440

```
Db 993 ----- 992
QY 1441 GCGGCCCCAACCTGGCTCTCTGGCCAGCTGCAGATCTACAGGGCATCTCTGAGGCC 1500
Db 993 ----- 992
QY 1501 AGAACCTGAGGTGGTGGGAGAGAGAGTTGTAGGCATGGAAGAGAGCCAGAGCCCC 1560
Db 993 ----- 992
QY 1561 GAAAGAGAGCCTGGGCCAGGCCACGATATAAACCCTCCGAGGGGTGATGAGTCCATCAG 1620
Db 993 ----- 992
QY 1621 TCTTCTGAGCCCTCTTGGAGCTGAGAGACGCTCAGAGACAGTGCAGATCCAGAGGT 1680
Db 993 ----- 992
QY 1681 CTTCTCTTCCAGAGATCTTCACATGAAAGAGCCCTGAGCCCTTCCACACAGCTTGCAAG 1740
Db 1037 CTTCTCTTCCAGAGATCTTCACATGAAAGAGCCCTGAGCCCTTCCACACAGCTTGCAAG 1096
QY 1741 GACCAAGGAGGCCAGCAGGTGGACAGAGGGGCTTACGCTGCTGGAAGTCCGCCAGTC 1800
Db 1097 GACCAAGGAGGCCAGCAGGTGGACAGAGGGGCTTACGCTGCTGGAAGTCCGCCAGTC 1156
QY 1801 AGTGTACCTCCAGAGGAGTGGCTGGTGGCCAAACCGGACCCAGGCCCTTCCAGAGGA 1860
Db 1157 AGTGTACCTCCAGAGGAGTGGCTGGTGGCCAAACCGGACCCAGGCCCTTCCAGAGGA 1216
QY 1861 GGAGCAGGGGCGAGGGGCGAGGGGCGAGAGAGCCCTGCAATTTCTCTACGCCAGGTTCCG 1920
Db 1217 GGAGCAGGGGCGAGGGGCGAGGGGCGAGAGAGCCCTGCAATTTCTCTACGCCAGGTTCCG 1276
QY 1921 GAAGTGTGTGAGACAGAGCCAGCTGCATGACAGTGGAGAGAGGGGCGAGGCCCTGAGCCCT 1980
Db 1277 GAAGTGTGTGAGACAGAGCCAGCTGCATGACAGTGGAGAGAGGGGCGAGGCCCTGAGCCCT 1336
QY 1981 CACAGATGGCCAGCGCTCCCTGACACTGAAGAGATCCCAACTCCTTGGAGAAACACCC 2040
Db 1337 CACAGATGGCCAGCGCTCCCTGACACTGAAGAGATCCCAACTCCTTGGAGAAACACCC 1396
QY 2041 TCAGCTGTGTTGCGCGACACATTTCTCAGCTCCGCCCCATACCCGCTACTACAGCCCTC 2100
Db 1397 TCAGCTGTGTTGCGCGACACATTTCTCAGCTCCGCCCCATACCCGCTACTACAGCCCTC 1456
QY 2101 ACCTCCACACCTGTCTACTACAGGCTCACCTCCACCCCTGTCTACTACAGCCCTCAC 2160
Db 1457 ACCTCCACACCTGTCTACTACAGGCTCACCTCCACCCCTGTCTACTACAGCCCTCAC 1516
QY 2161 TACAGCCTTAAGTCCAGGCCCCATGTGCTGCTCAAGGGGCTCAAGACTTCTTAAGTGG 2220
Db 1517 TACAGCCTTAAGTCCAGGCCCCATGTGCTGCTCAAGGGGCTCAAGACTTCTTAAGTGG 1576
QY 2221 GATGTGTAGAGGAGCTGAAGGTACCTTTGGGGGCAACAGCACCCTAGTTTCAATTCCTCA 2280
Db 1577 GATGTGTAGAGGAGCTGAAGGTACCTTTGGGGGCAACAGCACCCTAGTTTCAATTCCTCA 1636
QY 2281 CTCTAGCCCTGCACACTACCTGTGGCAGCGAATATAAAGAGAGCTTCCCGTCAAAAAG 2340
Db 1637 CTCTAGCCCTGCACACTACCTGTGGCAGCGAATATAAAGAGAGCTTCCCGTCAAAAAG 1696
QY 2341 GGTCAAGGCTCCACAGCCCGCCCTCTGCACTGCTGCTCTCTCCAGTTCATTTCTT 2400
Db 1697 GGTCAAGGCTCCACAGCCCGCCCTCTGCACTGCTGCTCTCTCCAGTTCATTTCTT 1756
QY 2401 GGAAACAGCAGGCGCCAGCAACCACTGGGCCCCCAAGAGGAGGAGAGATCTTACGCCCA 2460
Db 1757 GGAAACAGCAGGCGCCAGCAACCACTGGGCCCCCAAGAGGAGGAGAGATCTTACGCCCA 1816
QY 2461 GCGCGGGAGGCTGGAAGGCTGGCAGATCGCTTCCCTCATCCACCTCCACCGGTCCAGG 2520
Db 1817 GCGCGGGAGGCTGGAAGGCTGGCAGATCGCTTCCCTCATCCACCTCCACCGGTCCAGG 1876
```

```
QY 2521 TCTTTGCTGTGCTCCCAAGACCTCCTGTGACACACAGGCGAGATCAGAGGCCAGGCCA 2580
Db 1877 TCTTTGCTGTGCTCCCAAGACCTCCTGTGACACACAGGCGAGATCAGAGGCCAGGCCA 1936
QY 2581 GAGATAGTCTTCTTTTGTGCTTTCTGAGCCCTGTGAGTCAAGTTTTCATAGCTTACA 2640
Db 1937 GAGATAGTCTTCTTTTGTGCTTTCTGAGCCCTGTGAGTCAAGTTTTCATAGCTTACA 1996
QY 2641 GTATGTGCTTTGTACTGAGAAATATAACACATTTTCATATAAAAAAAAAAAAAA 2700
Db 1997 GTATGTGCTTTGTACTGAGAAATATAACACATTTTCATATTGTATTTCATAAAAAA 2056
QY 2701 AAAA 2704
Db 2057 AAAA 2060

RESULT 12
AAF63578
ID AAF63578 standard; cDNA: 1026 BP.
XX
AC AAF63578;
XX
DT 11-MAY-2001 (first entry)
XX
DE Human phosphatase NP_060746_h coding sequence.
XX
KW Phosphatase; gene therapy; enzyme; cancer; pathophysiological hypoxia;
KW cardiac dysfunction; vascular disorder; myopathy; ectodermal dysplasia;
KW congenital muscle disorder; Papillon-Lefevre syndrome; Cowden disease;
KW Moebius syndrome; Bjornstad syndrome; Bannayan Zonana syndrome;
KW schizophrenia; hamartoma; ss.
XX
OS Homo sapiens.
XX
PN WO200112819-A2.
XX
PD 22-FEB-2001.
XX
PF 11-AUG-2000; 2000WO-US22158.
XX
PR 13-AUG-1999; 99US-0149005.
XX
PA (SUGEN-) SUGEN INC.
XX
PI Plowman GD, Martinez R, Whyte D, Hill R, Flanagan P, Lioubin M;
XX
DR WPI: 2001-211226/21.
XX
DR P-PSDB: AAB73226.
XX
PT New protein phosphatase polypeptide for diagnosing and treating
PT phosphatase related disorders such as cancer, schizophrenia, cardiac
PT dysfunction and/or vascular disorders
XX
PS Claim 1; Fig 4; 138pp; English.
XX
CC The present invention relates to phosphatase proteins and coding
CC sequences. The present sequence is one such phosphatase coding sequence.
CC Phosphatases are enzymes that catalyse the dephosphorylation of proteins
CC modified by phosphorylation of serine, threonine or tyrosine residues.
CC The phosphatases are useful for treating a variety of diseases: for
CC example cancer e.g. breast, urogenital, prostate, head, neck, lung
CC cancers, synovial sarcomas, renal cell carcinoma, non-small cell lung
CC cancer, hepatocellular carcinoma, pancreatic endocrine tumours, stomach
CC cancer, glioblastoma, colorectal cancer and thyroid cancer,
CC pathophysiological hypoxia, cardiac dysfunction and/or vascular
CC disorders, myopathies, congenital muscle disorders, Papillon-Lefevre
CC syndrome, Cowden disease, ectodermal dysplasia, Moebius syndrome,
CC Bjornstad syndrome, Bannayan Zonana syndrome, schizophrenia and
CC hamartomas.
XX
SQ Sequence 1026 BP: 217 A; 324 C; 326 G; 159 T; 0 other;
```

Query Match 30.3%; Score 820; DB 22; Length 1026;
Best Local Similarity 90.4%; Pred. No. 6.5e-155;
Matches 928; Conservative 0; Mismatches 0; Indels 98; Gaps 1;

```
QY 1048 ATGCTGCTGCTGGTGGCAGACGCGGAGCCGCTCCCGCATCTTCCCGCACTCTACTG 1107
      |||||||
Db 1 ATGCTGCTGCTGGTGGCAGACGCGGAGCCGCTCCCGCATCTTCCCGCACTCTACTG 60
QY 1108 GGGTCAGAGTGAAGCAGCAAACTGGAGAGTGTGAGAGGAAGAGGATGACCCACATC 1167
      |||||||
Db 61 GGGTCAGAGTGAAGCAGCAAACTGGAGAGTGTGAGAGGAAGAGGATGACCCACATC 120
QY 1168 TTGAACATGGCCCGGAGATTGACAACTTACCTGAGGCTTACACTACCAATGTG 1227
      |||||||
Db 121 TTGAACATGGCCCGGAGATTGACAACTTACCTGAGGCTTACACTACCAATGTG 180
QY 1228 CGCCTCTGGATGAGAGTGGCCCACTGCTGCCCATCTGGAAGAGACGACCGCTTC 1287
      |||||||
Db 181 CGCCTCTGGATGAGAGTGGCCCACTGCTGCCCATCTGGAAGAGAGACGACCGCTTC 240
QY 1288 ATTGAGGCTGCAAGACAGACAGGCGACCGTGTGCTCCACTGCAAGATGGGCTGAGC 1347
      |||||||
Db 241 ATTGAGGCTGCAAGACAGACAGGCGACCGTGTGCTCCACTGCAAGATGGGCTGAGC 300
QY 1348 CGCTCAGCGGCCACAGTGTGCTGCTATGCCATGAGCAATGCAAGTACAGCTGGAGCAG 1407
      |||||||
Db 301 CGCTCAGCGGCCACAGTGTGCTGCTATGCCATGAGCAATGCAAGTACAGCTGGAGCAG 360
QY 1408 GCGCTGGCCGCGCAGGAGAGCTCGGGCCATGGCCCGCCCAACCTGGCTTCCTGCGC 1467
      |||||||
Db 361 GCGCTGGCCGCGCAGGAGAGCTCGGGCCATGGCCCGCCCAACCTGGCTTCCTGCGC 420
QY 1468 CAGCTGAGATCTACCAAGGCGATCCTGACG----- 1497
      |||||||
Db 421 CAGCTGAGATCTACCAAGGCGATCCTGAGGCCAGCCGCGCAAGCCATGTCGTGGAGCAG 480
QY 1498 ----- 1497
      |||||||
Db 481 AAAGTGGTGGGGTCTCCCGCAGAGAGAGCACCACCCCTGAAGTCTTACACATTCCCA 540
QY 1498 -----GCCAAGACCTGAGGCTGTGGGAGAGAGAAAGTTGAGCATGGAAGAGAGC 1549
      |||||||
Db 541 CCTCTTCCGCGAAGACCTGAGGCTGTGGGAGAGAAAGTTGAGCATGGAAGAGAGC 600
QY 1550 CAGCAGACCCCGAAAGAGAGCCTGGGCGACAGCGCCACGTATAAAGCTCCGAGGGGTATG 1609
      |||||||
Db 601 CAGCAGACCCCGAAAGAGAGCCTGGGCGACAGCGCCACGTATAAAGCTCCGAGGGGTATG 660
QY 1610 AGGTCCATCACTTCTTGAGACCTCCTTGAGAGCTGAGAGCACCTCAGAGACCACTGAC 1669
      |||||||
Db 661 AGGTCCATCACTTCTTGAGACCTCCTTGAGAGCTGAGAGCACCTCAGAGACCACTGAC 720
QY 1670 ATGCGAGAGGCTTCTCTCCAGAGAGCTTCAATGAAGAGCCCTGACAGCCCTTCCCA 1729
      |||||||
Db 721 ATGCGAGAGGCTTCTCTCCAGAGAGCTTCAATGAAGAGCCCTGACAGCCCTTCCCA 780
QY 1730 CAGCTTGAAGAGACCAAGGAGGAGCAGCAGTGAAGAGGGGCTCAGCTGACCTGAAG 1789
      |||||||
Db 781 CAGCTTGAAGAGACCAAGGAGGAGCAGCAGTGAAGAGGGGCTCAGCTGACCTGAAG 840
QY 1790 TCCCGCCAGTACGTGTTACCTCCAGGAGCAGTGCCTGGTGGCCAAACCGACAGGCC 1849
      |||||||
Db 841 TCCCGCCAGTACGTGTTACCTCCAGGAGCAGTGCCTGGTGGCCAAACCGACAGGCC 900
QY 1850 TTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1909
      |||||||
Db 901 TTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960
QY 1910 CCCAGGTTCCGAAAGGTGTGAGACAGGCGCATGATGATGAGTGAAGAGAGGAGGAGG 1969
      |||||||
Db 961 CCCAGGTTCCGAAAGGTGTGAGACAGGCGCATGATGATGAGTGAAGAGAGGAGGAGG 1020
```

QY 1970 GCCTGA 1975
Db 1021 GCCTGA 1026

RESULT 13

ABL40803
ID ABL40803 standard; DNA; 599 BP.

AC ABL40803;

DT 03-JUL-2002 (first entry)

XX Human MAP kinase phosphatase-like enzyme DNA fragment.

XX Mitogen activated protein; MAP; MAP kinase phosphatase-like enzyme;

KW antiparkinsonian; cerebrotrophic; neuroprotective; nootropic; gene;

KW antilepteric; anticonvulsant; anti-HIV; antiarrhythmic; hypotensive;

XX antiallergic; dermatological; vulnerary; gene therapy; ds.

OS Homo sapiens.

PN MO200220732-A2.

PD 14-MAR-2002.

PF 27-AUG-2001; 2001WO-EP09848.

PR 07-SEP-2000; 2000US-230709P.

PA (FARB) BAYER AG.

PI Liou J;

DR MPI; 2002-339802/37.

PT New human mitogen activated protein kinase phosphatase-like enzyme

PT polypeptide, regulators of which are useful for preventing, treating

PT allergies including asthma, diabetes, obesity, cancer and

PT cardiovascular diseases

PS Disclosure; Fig 8; 134pp; English.

XX The invention relates to a purified human mitogen activated protein (MAP)

XX kinase phosphatase-like enzyme polypeptide. The enzyme can be expressed

XX by standard recombinant methodology. The MAP kinase phosphatase-like

XX enzyme and encoding polynucleotides are useful for screening for

XX modulators which are used for treating a MAP kinase phosphatase-like

XX enzyme dysfunction related disease, such as asthma, a central nervous

XX system disorder, diabetes, obesity, chronic obstructive pulmonary

XX disease, cancer or a cardiovascular disease. The enzyme can be regulated

XX to treat allergies including asthma, allergic rhinitis, atopic

XX dermatitis, and anaplasia; central nervous system disorders such as

XX brain injuries, Parkinson's disease, dementia, multiple sclerosis, or

XX stroke, Alzheimer's disease, Huntington's disease, schizophrenia, Pick's

XX disease, Creutzfeldt-Jacob dementia, progressive nuclear palsy, and human

XX immunodeficiency virus (HIV) dementia, and cardiovascular diseases

XX including myocardial infarction, ischemic diseases of the heart, atrial

XX and ventricular arrhythmia, hypertensive vascular diseases and peripheral

XX vascular diseases. The enzyme is useful in diagnostic assays for

XX detecting diseases and abnormalities or susceptibility to diseases or

XX abnormalities related to the presence of mutations in the encoding

XX nucleic acid sequences. The present sequence represents the human MAP

XX kinase phosphatase-like enzyme DNA fragment.

SO Sequence 599 BP; 135 A; 185 C; 183 G; 96 T; 0 other;

Query Match 22.2%; Score 599; DB 24; Length 599;

Best Local Similarity 100.0%; Pred. No. 1.1e-110;

Matches 599; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1117 TGGACGCGACCAACTGGAGAGCTGACAGAGCAAGGCTCACCAATCTTGAACATG 1176

```

|||||
Db 1 TCGACGACGCAACCTGGAGAGCTGCAGAGAACAGGCTCCACCACATCTTGAACATG 60
Oy 1177 GCGCGGAGATTGACAACTTACCTGAGCGCTTCACTACCACAAATGGCGCTTCG 1236
Db 61 GCGCGGAGATTGACAACTTACCTGAGCGCTTCACTACCACAAATGGCGCTTCG 120
Oy 1237 GATGAGAGTGGCGCCAGCTGCTGCCGACTGGAAGAGACGACCGCTTCAATGAGCT 1296
Db 121 GATGAGAGTGGCGCCAGCTGCTGCCGACTGGAAGAGACGACCGCTTCAATGAGCT 180
Oy 1297 GCAAGAGCAGAGGCAACCCAGCTGCTGCTCACTGCAAGATGGCGCTCAGCCCTCAGCG 1356
Db 181 GCAAGAGCAGAGGCAACCCAGCTGCTGCTCACTGCAAGATGGCGCTCAGCG 240
Oy 1357 GCCAGAGTGGCGCTTATGCGCTGTAAGAGTACGATGAGCTGGAGAGCGCCCTGGCG 1416
Db 241 GCCAGAGTGGCGCTTATGCGCTGTAAGAGTACGATGAGCTGGAGAGCGCCCTGGCG 300
Oy 1417 CACGTGACAGAGCTCCGGCCATCCGCCGCCCAACCCCTGCTCTGCGCCAGCTGCAG 1476
Db 301 CACGTGACAGAGCTCCGGCCATCCGCCGCCCAACCCCTGCTCTGCGCCAGCTGCAG 360
Oy 1477 ATCTACACAGGCGATCTCTGACGCGCAAGAACCTGAGGGTGGTGGAGAGAGGTTGTAGG 1536
Db 361 ATCTACACAGGCGATCTCTGACGCGCAAGAACCTGAGGGTGGTGGAGAGAGGTTGTAGG 420
Oy 1537 CATGGAAGAGGACGCGGACCGCCGAAAGAGAGTACGCTGGCGCCAGCGCTATTAACCT 1596
Db 421 CATGGAAGAGGACGCGGACCGCCGAAAGAGAGTACGCTGGCGCCAGCGCTATTAACCT 480
Oy 1597 CCGAGGGGTCATGAGGTCATCAGTCTTCTGAGGCTCTCTTGAGAGTGGAGAGCACTC 1656
Db 481 CCGAGGGGTCATGAGGTCATCAGTCTTCTGAGGCTCTCTTGAGAGTGGAGAGCACTC 540
Oy 1657 AGAGACCACTGACATGCGCAGAGAGGTTCTTCTCCACAGAGTCTTACATGAAGACCTC 1715
Db 541 AGAGACCACTGACATGCGCAGAGAGGTTCTTCTCCACAGAGTCTTACATGAAGACCTC 599

```

RESULT 14

```

ABLA0802
ID ABLA0802 standard; DNA; 409 BP.
AC ABLA0802;
XX
DT 03-JUL-2002 (first entry)
XX
DE Human MAP kinase phosphatase-like enzyme DNA fragment.
XX
KW Mitogen activated protein; MAP; MAP kinase phosphatase-like enzyme;
KW antidiabetic; anorectic; cytosolic; cardiant; human;
KW antiparkinsonian; cerebroprotective; neuroprotective; neurotropic; gene;
KW neuroleptic; anticonvulsant; anti-HIV; antirhythmic; hypotensive;
KW antiallergic; dermatological; vulnerary; gene therapy; ds.
XX
OS Homo sapiens.
XX
PN WO2002/20732-A2.
XX
PD 14-MAR-2002.
XX
PF 27-AUG-2001; 2001WO-EP09848.
XX
PR 07-SEP-2000; 2000US-230709P.
XX
PA (FARB ) BAYER AG.
XX
PI Liou J.
XX
DR WPI: 2002-339802/37.
XX
PT New human mitogen activated protein kinase phosphatase-like enzyme

```

```

PT polypeptide, regulators of which are useful for preventing, treating
PT allergies including asthma, diabetes, obesity, cancer and
PT cardiovascular diseases
PS Disclosure: Fig 7; 134pp; English.
XX
CC The invention relates to a purified human mitogen activated protein (MAP)
CC kinase phosphatase-like enzyme polypeptide. The enzyme can be expressed
CC by standard recombinant methodology. The MAP kinase phosphatase-like
CC enzyme and encoding polynucleotides are useful for screening for
CC modulators which are used for treating a MAP kinase phosphatase-like
CC enzyme dysfunction related disease, such as asthma, a central nervous
CC system disorder, diabetes, obesity, chronic obstructive pulmonary
CC disease, cancer or a cardiovascular disease. The enzyme can be regulated
CC to treat allergies including asthma, allergic rhinitis, atopic
CC dermatitis, and anaphylaxis, central nervous system disorders such as
CC brain injuries, Parkinson's disease, dementia, multiple sclerosis,
CC stroke, Alzheimer's disease, Huntington's disease, schizophrenia, Pick's
CC disease, Creutzfeldt-Jacob dementia, progressive nuclear palsy, and human
CC immunodeficiency virus (HIV) dementia, and cardiovascular diseases
CC including myocardial infarction, ischemic diseases of the heart, atrial
CC and ventricular arrhythmia, hypertensive vascular diseases and peripheral
CC vascular diseases. The enzyme is useful in diagnostic assays for
CC detecting diseases and abnormalities or susceptibility to diseases or
CC abnormalities related to the presence of mutations in the encoding
CC nucleic acid sequences. The present sequence represents the human MAP
CC kinase phosphatase-like enzyme DNA fragment.

```

Sequence 409 BP; 87 A; 127 C; 121 G; 73 T; 1 other:

Query Match 14.7%; Score 397; DB 24; Length 409;
 Best Local Similarity 99.5%; Pred. No. 3e-70;
 Matches 408; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```

Oy 912 GCGATCCGCTGCTGACGCTGGAAGTGTGATGTCAGTACCTGAGAGTGCATCTC 971
Db 1 GCGATCCGCTGCT-AGCTGTGGAAGTGTGATGTCAGTACCTGAGAGTGCATCTC 59
Oy 972 CAAAGAGATCCGCGCAGGCTGAGAGTGGCGCTGGGGCTCCCGCCAGCATACCGTGA 1031
Db 60 CAAAGAGATCCGCGCAGGCTGAGAGTGGCGCTGGGGCTCCCGCCAGCATACCGTGA 119
Oy 1032 CTTTCATGCAACAACGATGCTGCTCTGTGACACAGCGGAGCGCTCCGCTCTT 1091
Db 120 CTTTCATGCAACAACGATGCTGCTCTGTGACACAGCGGAGCGAGCTCCGCTCTT 179
Oy 1092 CCCCCACTCTTACCTGGGCTCAGAGTGAAGCAGCAACCTGAGAGCTGCAGAGGA 1151
Db 180 CCCCCACTCTTACCTGGGCTCAGAGTGAAGCAGCAACCTGAGAGCTGCAGAGGA 239
Oy 1152 CAGGGTCAACCCATCTGAACATGCGCCGGAGATGACAACTTCACTCCGAGCGCTT 1211
Db 240 CAGGGTCAACCCATCTGAACATGCGCCGGAGATGACAACTTCACTCCGAGCGCTT 299
Oy 1212 CACCTACCAATGTGCCCTCTGGATGAGAGTGGCGCCAGCTGTCGCGACTGGAA 1271
Db 300 CACCTACCAATGTGCCCTCTGGATGAGAGTGGCGCCAGCTGTCGCGACTGGAA 359
Oy 1272 GGAGAGCAGCGCTTCAATGAGGCTGCAAGAGACAGGCGACCCAGCTGC 1321
Db 360 GGAGAGCAGCGCTTCAATGAGGCTGCAAGAGACAGGCGACCCAGCTGC 409

```

RESULT 15

```

AAH98183/C
ID AAH98183 standard; cDNA; 717 BP.
AC AAH98183;
XX
DT 12-OCT-2001 (first entry)
XX
DE Human EST-derived coding sequence SEQ ID NO: 40.
XX

```


OY	1441	GGCGGCCCCAACCTGGCTTCCTCGGCGCCAGCTGCGAGATCTACAGAGGATCTCTACG---	1497
Db	1444	GGCGGCCCCAACCTGGCTTCCTCGGCGCCAGCTGCGAGATCTACAGAGGATCTCTACGCGCC	1503
OY	1498	-----	1497
Db	1504	AGCGCCGAGGCCATGTCTGGGAGCAGAAATGSGGTGGGTCCTCCAGAGAGCACCCA	1563
OY	1498	-----GCCAAGAACCGAGAGGTGGGGAG	1522
Db	1564	GGCCCTGAAGTCTCTACACCATTCCACCTCTTCCGCCAAGACCTGAGAGGTGGGGAG	1623
OY	1523	GAGAGGTTTGAAGCATGGAAGAGACCGAGCAGCCCGGAAAGAAAGAGCCTGGGCCAAG	1582
Db	1624	GAGAAAGTTTGAAGCATGGAAGAGACCGAGCAGCCCGGAAAGAAAGAGCCTGGGGCAAG	1683
OY	1583	CCACGTATAAACCCTCGAGGGGTCATGAGGTGCATTCAGTCTTCTGAGACCTCTTGGAG	1642
Db	1684	CCACGTATAAACCCTCGAGGGGTCATGAGGTGCATTCAGTCTTCTGAGACCTCTTGGAG	1743
OY	1643	CTGAGAGGACACTCAGAGACCAAGTGCATGACGAGAGGTCTTCTCTCCAGAGATCTTCA	1702
Db	1744	CTGAGAGGACACTCAGAGACCAAGTGCATGACGAGAGGTCTTCTCTCCAGAGATCTTCA	1803
OY	1703	CATGAGAGCCTCTGCAGCCCTTCCACAGCTTGCAGAGACCAAGGAGGCCAGAGTG	1762
Db	1804	CATGAAAGCCTCTGCAGCCCTTCCACAGCTTGCAGAGACCAAGGAGGCCAGAGTG	1863
OY	1763	GACAGGGGGGCTCAGGCTGGCGGAAAGTCCGGCCAGTGCAGAGGTATACCTCAGGGGAGT	1822
Db	1864	GACAGGGGGGCTCAGGCTGGCGGAAAGTCCGGCCAGTGCAGAGGTATACCTCAGGGGAGT	1923
OY	1823	GCCCTGGTGGCCAAACCGAGCCAGGCTTCCAGAGACAGAGACAGGGGACAGGG	1882
Db	1924	GCCCTGGTGGCCAAACCGAGCCAGGCTTCCAGAGACAGAGACAGGGGACAGGG	1983
OY	1883	CAGGGAAGCCCTTCGATTTCTCTACGCCAGGTTCCGGAAGGTGTAGACAGGCCAGC	1942
Db	1984	CAGGGAAGCCCTTCGATTTCTCTACGCCAGGTTCCGGAAGGTGTAGACAGGCCAGC	2043
OY	1943	GTCGATACAGTGAAGAGAGGGGCGAGGCTCAGGCCCTACATGACCCAGCCTCCCTG	2002
Db	2044	GTCGATACAGTGAAGAGAGGGGCGAGGCTCAGGCCCTCAGACATGACCCAGCCTCCCTG	2103
OY	2003	ACACTGAAGAGATCCACAACCTCTTGGAAACACACCTCAGCTGTGTCGCGACACAT	2062
Db	2104	ACACTGAAGAGAGATCCACAACCTCTTGGAAACACACCTCAGCTGTGTCGCGACACAT	2163
OY	2063	TTCGCTAGATCGGCGCCCAATCCCGCATCTACAGGCTCAGCCCGGCTGCTACAG	2122
Db	2164	TTCGCTAGATCGGCGCCCAATCCCGCATCTACAGGCTCAGCCCGGCTGCTACAG	2223
OY	2123	GCCTCACCTCCACACCCTGTCACTACAGGCTCAGCTCTACAGCCTTAAGTCCAGGCC	2182
Db	2224	GCCTCACCTCCACACCCTGTCACTACAGGCTCAGCTCTACAGCCTTAAGTCCAGGCC	2283
OY	2183	ATGTCGCTGCTCCAGAGGCTCAAGCTTTCTTACGTGGATGTGGTAAGGAGACTGAAG	2242
Db	2284	ATGTCGCTGCTCCAGAGGCTCAAGCTTTCTTACGTGGATGTGGTAAGGAGACTGAAG	2343
OY	2243	TACCTTTGGGGGAAACAGCACCTTAATTTCATTTCAACTGACGCTCGACACTCACT	2302
Db	2344	TACCTTTGGGGGAAACAGCACCTTAATTTCATTTCAACTGACGCTCGACACTCACT	2403
OY	2303	GTCGACAGGAATGAAGACAGAGCTTCCGCTGCAAAAAGGGTCAGGCTTCCACCCCGCC	2362
Db	2404	GTCGACAGGAATGAAGACAGAGCTTCCGCTGCAAAAAGGGTCAGGCTTCCACCCCGCC	2463
OY	2363	CCCTCCCTGACCGCTGCTCTCTCCAGTTCATTCCTGGAACAGGCGAGGCGAGGCAAC	2422
Db	2464	CCCTCCCTGACCGCTGCTCTCTCCAGTTCATTCCTGGAACAGGCGAGGCGAGGCAAC	2523

QY	2423	CAGTGGCCCCAAAGGAGGAGGATCTTCAGAGGCCCGGAGAGCTGGAAGGCT	2482
Db	2524	CAGTGGCCCCAAAGGAGGAGGATCTTCAGAGGCCCGGAGAGCTGGAAGGCT	2583
QY	2483	GGCAGATGCTTCCCTCATTCACACCTCCACCGGTCAGGCTTTTGCTGTGCCAGACC	2542
Db	2584	GGCAGATGCTTCCCTCATTCACACCTCCACCGGTCAGGCTTTTGCTGTGCCAGACC	2643
QY	2543	TTCCTGTACACACGAGGATTCACAGAGGACACAGGCGCAGAGATATCTTTCTTTTGTCT	2602
Db	2644	TCTCTGTACACACGAGGATTCACAGAGGACACAGGCGCAGAGATATCTTTCTTTTGTCT	2703
QY	2603	TTCTGTGCTTGTGGCTAGTACAGTCTTTTCAATAGCCTTACAGTATCTGCTTTGTACTGAGAA	2662
Db	2704	TTCTGTGCTTGTGGCTAGTACAGTCTTTTCAATAGCCTTACAGTATCTGCTTTGTACTGAGAA	2763
QY	2663	ATTAACACATTTTCATATATAAAAAAAAAAAAAAAAAAAAAA	2704
Db	2764	ATTAACACATTTTCATATATTTGGTTAAAAAAAAAAAAAAAAAAAAA	2805
RESULT 2			
AX099939			
LOCUS	AX099939	2781 bp	DNA linear PAT 02-APR-2001
DEFINITION	Sequence 21 from Patent WO0120004.		
ACCESSION	AX099939		
VERSION	AX099939.1	GI:13538949	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 2781)		
AUTHORS	Yue,H., Tang,Y.T., Bandman,O., Hillman,J.L., Baughn,M.R., Azimzal,Y. and Lu,D.A.		
TITLE	Protein phosphatase and kinase proteins		
JOURNAL	Patent: WO 0120004-A 21-22-MAR-2001; Incyte Genomics, Inc. (US)		
FEATURES	location/Qualifiers		
source	1..2781		
BASE COUNT	576 a 906 c 820 g 479 t		
ORIGIN			
Query Match	94.8%; Score 2562.8; DB 6; Length 2781;		
Best Local Similarity	96.4%; Pred. No. 0;		
Matches 2672:	Conservative 0; Mismatches 2; Indels 98; Gaps 1;		
QY	17	GGGGTCCAGGACTCTCCGGGGGTTGAAGGAGGGCGTCCCGGTCAGCCAGCAGGT	76
Db	10	GGGGTCCAGGACTCTCCGGGGGTTGAAGGAGGGCGTCCCGGTCAGCCAGCAGGT	69
QY	77	GCTCGGGCTGTGCTCATATGCGCTGTGTACAGTAGCCGTTCCGCCCGGGAGCGGCG	136
Db	70	GCTCGGGCTGTGCTCATATGCGCTGTGTACAGTAGCCGTTCCGCCCGGGAGCGGCG	129
QY	137	CCTCAGCCCGCTGGGGCGCTGGGACCAAGCGCGTCACAGGAAGTAGTGCATCCAGCGAA	196
Db	130	CCTCAGCCCGCTGGGGCGCTGGGACCAAGCGCGTCACAGGAAGTAGTGCATCCAGCGAA	189
QY	197	GGCAGAGCTTTCGCGTGTCCGTCGGGCGTGTCTGGACTGCAGAGATGGAGGGAGCAATG	256
Db	190	GGCAGAGCTTTCGCGTGTCCGTCGGGCGTGTCTGGACTGCAGAGATGGAGGGAGCAATG	249
QY	257	ATGATTCAGACAGAGCCAGTTCGTAGCCAAACAGAGAGAGGCCCGGATGAGAGAGCTCC	316
Db	250	ATGATTCAGACAGAGCCAGTTCGTAGCCAAACAGAGAGAGGCCCGGATGAGAGAGCTCC	309
QY	317	ACGGGACCAAGACACTTCCGGCAAGATGCCAAGATCCCAAGAAAGCAGAGGAGACAGA	376

Oy	377	GGCAGACCTGCACTCATGTGACAGCTGCTAAGGCCGAGATGACATCCGCTGGCAG	436
Db	370	GGCACCACCTGTCACTCATGTGACAGCTGCTAAGGCCGAGATGACATCCGCTGGCAG	429
Oy	437	CCGACCTGGAGCACCCCGGCTCCCGGCTCCGGTACCTGCTGGAGTTTCTACACGAG	496
Db	430	CCGACCTGGAGCACCCCGGCTCCCGGCTCCGGTACCTGCTGGAGTTTCTACACGAG	489
Oy	497	AAGGAGAAAGTCTGAGCCAGATGAGACGGTCTCTGGGCGTGATTTTCCCTGACAGA	556
Db	490	AAGGAGAAAGTCTGAGCCAGATGAGACGGTCTCTGGGCGTGATTTTCCCTGACAGA	549
Oy	557	GGTCCCCCAGCTGCACCCCGGCGTGGTCTTGCCCCCTTGGAGTGACACCCAGGTCTAC	616
Db	550	GGTCCCCCAGCTGCACCCCGGCGTGGTCTTGCCCCCTTGGAGTGACACCCAGGTCTACT	609
Oy	617	TGATGAGAACGGGGGCTTCAGGCTGACCTGTGGGGGCAAAAGCCGGATCTTCAAGCCCA	676
Db	610	TGATGAGAACGGGGGCTTCAGGCTGACCTGTGGGGGCAAAAGCCGGATCTTCAAGCCCA	669
Oy	677	TCTCCATCCAGACATGTTGGGCGCACACTCCAGGTATTGCACCAAGATGTGAGGACGATC	736
Db	670	TCTCCATCCAGACATGTTGGGCGCACACTCCAGGTATTGCACCAAGATGTGAGGACGATC	729
Oy	737	TAGGACGGGGCTTTGTACCGGGGTGGAGTGGCCCTCACTGGGCGACCACTACAGAGAGA	796
Db	730	TAGGACGGGGCTTTGTACCGGGGTGGAGTGGCCCTCACTGGGCGACCACTACAGAGAGA	789
Oy	797	GACTGAACCTCCGACAGAGACTGGCTCAATATGATGAGAGCGGTATGGCGGACCTGGAGCTTC	856
Db	790	GACTGAACCTCCGACAGAGACTGGCTCAATATGATGAGAGCGGTATGGCGGACCTGGAGCTTC	849
Oy	857	TGGGGGCTCCACAGCGCGGAGCCCTGGGCGGTCTCAGAACAGAGAGATGAGACAGGCA	916
Db	850	TGGGGGCTCCACAGTGCCGAGCCTGGGCGGCTCCTCAGAACAGAGAGATGAGACAGGCA	909
Oy	917	TCCGCTGCTGAGCTGTGGAAAGTTTGGATGTCAGTGACCTGGAGAGTGTCACTTCCAAAG	976
Db	910	TCCGCTGCTGAGCTGTGGAAAGTTTGGATGTCAGTGACCTGGAGAGTGTCACTTCCAAAG	969
Oy	977	AGATTCGGCGGCTGTGGAGACTCGCGCTGGGGGCTTCCCGCTCCAGAGTACCGTGACTTCA	1036
Db	970	AGATTCGGCGGCTGTGGAGACTCGCGCTGGGGGCTTCCCGCTCCAGAGTACCGTGACTTCA	1029
Oy	1037	TCGACAACAGATGCTGTGCTGTGTGGCACAGCGGAGCCGAGCCTCCCGCATTTCCCCC	1096
Db	1030	TCGACAACAGATGCTGTGCTGTGTGGCACAGCGGAGCCGAGCCTCCCGCATTTCCCCC	1089
Oy	1097	ACCTCAACCGGGGCTCAGATGGAAGCCACAAACCTGGAGAGAGCTCAGAGGAAACAGGG	1156
Db	1090	ACCTCAACCGGGGCTCAGATGGAAGCCACAAACCTGGAGAGAGCTCAGAGGAAACAGGG	1149
Oy	1157	TCACCCACATCTTGAACATGGGCCGGGAGATTGACAATTTCAACCCTGAGCGCTTCACCT	1216
Db	1150	TCACCCACATCTTGAACATGGGCCGGGAGATTGACAATTTCAACCCTGAGCGCTTCACCT	1209
Oy	1217	ACCACAATGTGGCGCTCTGGGATGAGAGATGTCGGCCACGCTGTCCGCACTGGAGAGAGA	1276
Db	1210	ACCACAATGTGGCGCTCTGGGATGAGAGATGTCGGCCACGCTGTCCGCACTGGAGAGAGA	1269
Oy	1277	CGCACCGCTTCATTGAGGCTGCAAGAGACACAGGGCACCCACGCTGCTGGTCCACTGCAGAGA	1336
Db	1270	CGCACCGCTTCATTGAGGCTGCAAGAGACACAGGGCACCCACGCTGCTGGTCCACTGCAGAGA	1329
Oy	1337	TGGGCGTCAACCGCTCAAGCGGCCACAGTGTGGGCTATGCCATGAGAACAGTACGATGCA	1396
Db	1330	TGGGCGTCAACCGCTCAAGCGGCCACAGTGTGGGCTATGCCATGAGAACAGTACGATGCA	1389
Oy	1397	GCCCTGGAGAGAGCCCTGCGGCACGTGAGAGAGCTTCGGGCCCATGCGCCGCCCAAAACCTG	1456
Db	1390	GCCCTGGAGAGAGCCCTGCGGCACGTGAGAGAGCTTCGGGCCCATGCGCCGCCCAAAACCTG	1449

QY	1457	GCCTCTCGCGCAGCTGCGAGATCTACGAGGCATCTCTACG-----	1497
Db	1450	GCCTCTCGCGCAGCTGCGAGATCTACGAGGCATCTCTACG-----	1509
QY	1498	-----	1497
Db	1510	TCCTGGAGACAGAAAGTGGGTGGGGTCTCCCGAGAGAGACCACGCCCTGAASTCTCTA	1569
QY	1498	-----GCCAGAACCTGAGGGTGGTGGGGAGAGAGGTTGTAGGCA	1538
Db	1570	CACCATTCGCCACCTCTCCGCCAGAACCTGAGGGTGGTGGAGAGAGAGGTTGTAGGCA	1629
QY	1539	TGGAGAGAGCCGAGGACAGCCCCCAAGAGAGCCCTGGGCAAGGCCACGTATTAACCTCC	1598
Db	1630	TGGAGAGAGCCGAGGACAGCCCCCAAGAGAGCCCTGGGCAAGGCCACGTATTAACCTCC	1689
QY	1599	GAGGGCTCATGAGGTCCATCAGTCTTCTGAGCCCTCCTTGGAGTGGAGAGCACTCAG	1658
Db	1690	GAGGGCTCATGAGGTCCATCAGTCTTCTGAGCCCTCCTTGGAGTGGAGAGCACTCAG	1749
QY	1659	AGACCACTGAGATCCGAGAGTCTTCTCTCCAGAGTCTTCACATGAAAGCCTCTGC	1718
Db	1750	AGACCACTGAGATCCGAGAGTCTTCTCTCCAGAGTCTTCACATGAAAGCCTCTGC	1809
QY	1719	AGCCCTTCCCACTGTCGAAGACCAAGAGAGCCAGAGSTGGACAGGGGGGCTCAGC	1778
Db	1810	AGCCCTTCCCACTGTCGAAGACCAAGAGAGCCAGAGSTGGACAGGGGGGCTCAGC	1869
QY	1779	CTGGCCTCAAGTCCCGGCAGTCAAGTGGTTACCTCCAGGGCAGTCCCGTGGTGGCCAAAC	1838
Db	1870	CTGGCCTCAAGTCCCGGCAGTCAAGTGGTTACCTCCAGGGCAGTCCCGTGGTGGCCAAAC	1929
QY	1839	GGAGCCAGGCTTCCAGAGACGAGAGCCAGGGGAGGGGAGGGGAGAGCCCTGCA	1898
Db	1930	GGAGCCAGGCTTCCAGAGAGAGAGAGCGGGGAGGGGAGGGGAGAGAGCCCTGCA	1989
QY	1899	TTTCTCTACGCCCAAGTCTCCGGAAGTGGTGAAGAGCCAGCGTGGATGACAGTGGAG	1958
Db	1990	TTTCTCTACGCCCAAGTCTCCGGAAGTGGTGAAGAGCCAGCGTGGATGACAGTGGAG	2049
QY	1959	AGGAGGGGAGAGCCCTGAGACCCTGCACATGCCAGCCTCCCTGACACTGAAGAGATCC	2018
Db	2050	AGGAGGGGAGAGCCCTGAGACCCTGCACATGCCAGCCTCCCTGACACTGAAGAGATCC	2109
QY	2019	ACAACCTCCTTGGAGAAACACCCCTCAGCTGTGTGCCGACACATTCCTCTCAGCTCGCC	2078
Db	2110	ACAACCTCCTTGGAGAAACACCCCTCAGCTGTGTGCCGACACATTCCTCTCAGCTCGCC	2169
QY	2079	CCATACCCGTCATGACAGCTCAGCCCTCCACCCCTGTCACTAGGGGCTCAGCTCCACCC	2138
Db	2170	CCATACCCGTCATGACAGCTCAGCCCTCCACCCCTGTCACTAGGGGCTCAGCTCCACCC	2229
QY	2139	CTGTACCTACACCCCTCAGCTCTGACAGCCTTAAGTCCAGAGGCCACATGCTGCCTGCCAA	2198
Db	2230	CTGTACCTACACCCCTCAGCTCTGACAGCCTTAAGTCCAGAGGCCACATGCTGCCTGCCAA	2289
QY	2199	GGGCTCAAGACTTTTAACTGGGATGTGTAGAAGGACTGAAGGTAACCTTTGGGGGCAAC	2258
Db	2290	GGGCTCAAGACTTTTAACTGGGATGTGTAGAAGGACTGAAGGTAACCTTTGGGGGCAAC	2349
QY	2259	AGCACCCTAGTTTCACTTCTCAACTGTAGCCCTGCACACTCAGCTGTGGGAGGGAATGAAA	2318
Db	2350	AGCACCCTAGTTTCACTTCTCAACTGTAGCCCTGCACACTCAGCTGTGGGAGGGAATGAAA	2409
QY	2319	ACAGACCTTCCGTCGAAAAAGGGGCAAGGCTCCACACCCCGCCCTCCCTGCACTGCC	2378
Db	2410	ACAGACCTTCCGTCGAAAAAGGGGCAAGGCTCCACACCCCGCCCTCCCTGCACTGCC	2469
QY	2379	TGTCTCTCTCCAGTTTCATCTTCGAGACAGAGCCAGGCAACAGTGGCCCCCAAGG	2438
Db	2470	TGTCTCTCTCCAGTTTCATCTTCGAGACAGAGCCAGGCAACAGTGGCCCCCAAGG	2529
QY	2439	CAGGACAGATCTCTCAGGCCCAAGCCGCGGAGAGCTGGAAGGCTGGCAGATGCTTCCCT	2498

|||||
Db 2530 CAGCAGATCTCAGGCGCCAGCGGCGGAGGCTGGAAGGCTGGCAGATGCTTCCCT 2589
Oy 2499 CATCAGCTCCAGCGGTCAGAGTCTTGTCTGCTCCCACTCTCTGACACACAGC 2538
Db 2590 CATCCACTCCACGGGTCAGAGTCTTGTCTCTCCCACTCTCTGACACACAGC 2649
Oy 2559 CAGATCAGAGGACACAGGAGAGATAGTCTTGTCTCTTGTGCGCTCTGGCTA 2618
Db 2650 CAGATCAGAGGACACAGGAGAGATAGTCTTGTCTCTTGTGCGCTCTGGCTA 2709
Oy 2619 GTCAGTTTTCATAGCTTACAGTATCTGCTTGTACTGAGAAATAACACATTTTCA 2678
Db 2710 GTCAGTTTTCATAGCTTACAGTATCTGCTTGTACTGAGAAATAACACATTTTCA 2769
Oy 2679 TAAAAAATAA 2690
Db 2770 TAAAAAATAA 2781
RESULT 3
AK094226
LOCUS AK094226 2905 bp mRNA linear PRI 15-JUL-2002
DEFINITION Homo sapiens cDNA FLJ36907 fis, clone BRACE2003800, weakly similar
to MAP kinase phosphatase.
ACCESSION AK094226
VERSION AK094226.1 GI:21753246
KEYWORDS oligo capping, fis (full insert sequence).
SOURCE Homo sapiens cerebellum cDNA to mRNA, clone_lib:BRACE2
clone:BRACE2003800.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Tanigami,A., Fujiwara,T., Shibahara,T., Goto,Y., Hiraio,M.,
Shimizu,F., Wakebe,H., Ono,T., Hishigaki,H., Watanabe,T., Ozaki,K.,
Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S.,
Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T.,
Kamura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M.,
Kikuchi,H., Kanda,K., Magatsuma,M., Murakawa,K., Kanehori,K.,
Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y.,
Sugano,S., Nagahara,K., Masuno,Y., Nagai,K. and Isogai,T.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 2905)
Isogai,T. and Yamamoto,J.
Direct Submission
Submitted (04-JUL-2002) Takao Isogai, FLJ Project(HRI Team): 2-6-7
Kazusa-Kamatairi, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing;
Research Association for Biotechnology (RAB): cDNA library
construction; Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'- and one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
FEATURES
Source
1: 2905
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="BRACE2003800"
/tissue_type="cerebellum"
/clone_lib="BRACE2"
/note="cloning vector: pME185FL3"
646..2187
/note="unnamed protein product"
/codon_start=1
/protein_id="BAC04314.1"
/db_xref="GI:21753247"
/translation="MAPKIPSPASSPCTGLVLPMSDVTQVYLDGDSGFSVTGGGOSR
IFPISITQTMATLQVLRHQAACEALGSLVPGSALTWASHYQERLNSQSCINEMTA

MADLESLRPPSABEGSSSEOEOMEQAIIRAEIMKVLVDYSLETSKIEIRQALERLGL
PLOOYRDFIDNOMILVADRSARSLIPPHLYLSENNAAEELQNRVTHINMARE
IDNFPERFTYHNVRILWDESSAOLIPMKETHRIIEAAROGHTVIVHCKMGVSRSA
TVLAVANKOYECESLEQALRHVQELRPARINRPFELROJYQIGILTRASRSHWEKV
GGVSEBHPAEVSTPPLPPEBEGGGRVVMGBESQAPAEERPRRLIDRGVM
RSISLPLESLESTSETSDMPFVSSHESHEPDPOLPQLATKGGQDVDPQPA
LKSROSVVTLQGSVAVANRQAFQEOBQOGOGGQEGPCISSTPFRKRVVQASVHDSG
EEGBA"
BASE COUNT 591 a 940 c 867 g 507 t
ORIGIN
Query Match 87.0%; Score 2352.8; DB 9; Length 2905;
Best Local Similarity 91.3%; Pred. No. 0;
Matches 2644; Conservative 0; Mismatches 2; Indels 250; Gaps 4;
Oy 31 GTCCGGGGGTTGAGAGGAGGAGGCGCCGCTGTCAGGCCAGTGTCTCGGCGCTGC 90
Db 1 GTCCGGGGGTTGAGAGGAGGAGGCGCCGCTGTCAGGCCAGTGTCTCGGCGCTGC 60
Oy 91 TCCATAGCCCTGTGACACAGTACAGTACGCGTCCGCCGCGGACAGCGCGCTGC 150
Db 61 TCCATAGCCCTGTGACACAGTACAGTACGCGTCCGCCGCGGACAGCGCGCTGC 120
Oy 151 GGGCCCTG----- 158
Db 121 GGGCCCTGGAATTCCTGAGAGAGGAGGAGACAGCCCTCCCGGCTCACCTGCGCTG 180
Oy 159 -----GGACAGGCGGCTCCAGAGAGAGTCCAGTCCAGAGAGCAG 201
Db 181 GGGCTCTCTCTCGGACAGGAGGAGGCGCTCCAGAGAGAGTCTGACCTCAGCAG 236
Oy 202 AGCTTTGGGCTGTCGCCGAGGAGGAGGCGTCTGCGACTGACAGTGGAGGAGCAATGATGAT 261
Db 237 AGTTTGGGCTGTCGCCGAGGAGGAGGCGTCTGCGACTGACAGTGGAGGAGCAATGATGAT 296
Oy 262 GCACAGAGGCGCACTTCTGAGCCCAACAGAGAGGCGCCGAGTGAAGAGAGCTCCACGCG 321
Db 297 GCACAGAGGCGCACTTCTGAGCCCAACAGAGAGGCGCCGAGTGAAGAGAGCTCCACGCG 356
Oy 322 GACCAGAGAGCTTCGGGCAAGAGATCCCAAGTCCCAAGAGAGAGAGAGAGAGAGAG 381
Db 357 GACCAGAGAGCTTCGGGCAAGAGATCCCAAGTCCCAAGAGAGAGAGAGAGAGAGAG 416
Oy 382 CACCTGCACCTCATGATGATACAGTCTGAGGCGGAGAGATGATGATCCGCTGGAGCCAG 441
Db 417 CACCTGCACCTCATGATGATACAGTCTGAGGCGGAGAGATGATGATCCGCTGGAGCCAG 476
Oy 442 CTGAGAGGACCCCGGCTCCCGGCTCCGCTACCTGCTGTAGTTTCTACACGAGAGAGA 501
Db 477 CTGAGAGGACCCCGGCTCCCGGCTCCGCTACCTGCTGTAGTTTCTACACGAGAGAGA 536
Oy 502 GAAGTCTGAGAGCAGATGAGAGAGGCTCTCTGGGCGCTGTGATTTCCCTGCAAG----- 554
Db 537 GAAGTCTGAGAGCAGATGAGAGAGGCTCTCTGGGCGCTGTGATTTCCCTGCAAGAGTTT 596
Oy 555 ----- 554
Db 597 GAGCAGGAGAGAGAGAGAGGAGGAGAGAGTGAAGAGAGAGGAGGAGCAATGCTTTCC 656
Oy 555 -----CAGCTCCCGCAGCTGACACCTGAGGCTGTGCTTTGCCCTCTGTGAGTGA 602
Db 657 TCTGTCCCTGCTGACAGCTCCCGCAGCTGACACCTGAGGCTGTGCTTTGCCCTCTGTGAGTGA 716
Oy 603 CACCCAGGTGATGATGATGAGAGAGGAGGCTTACAGTGTGAGAGCTGTGAGGAGAGAG 662
Db 717 CACCCAGGTGATGATGATGAGAGAGGAGGCTTACAGTGTGAGAGCTGTGAGGAGAGAG 776
Oy 663 GATCTTCAAGCCATCTTCCATCAGAGCATGTGGGCCACACTCCAGGATTTGACACAGC 722
Db 777 GATCTTCAAGCCATCTTCCATCAGAGCATGTGGGCCACACTCCAGGATTTGACACAGC 836
Oy 723 ATGTGAGGAGAGCTTGAAGGAGGAGGCTGTGATCCGAGGTGAGTGCCTCACCTGGAGCCAG 782
|||||

D	b	837	ATGTGAGGCAAGCTCTTAGGACGAGGGGCTTGTACCGGGTGGCAGTGGCCCTACACTGTGGGCGAG	896
Q	y	783	CCACTACGAGAGAGACTGAATCCGACAGAGCTGTCCTAATGATGTGACGGCTATGGC	842
D	b	897	CCACTACGAGAGAGACTGAATCCGACAGAGCTGTCCTAATGATGTGACGGCTATGAGC	956
Q	y	843	CGACCTGGAATCTCTGGGGCTCCCGAGGCGCAGGCTGGGGGGTCTCAGAAACGAGGAGCA	902
D	b	957	CGACCTGGAATCTCTGGGGCTCCCGAGGCGCAGGCTGGGGGGTCTCAGAAACGAGGAGCA	1016
Q	y	903	GATGGACGACGCGCATCCGTGTAGCTGTGGAAAGTGTGTGATGTCACTGACTGTGAGAG	962
D	b	1017	GATGGACGACGCGCATCCGTGTAGCTGTGGAAAGTGTGTGATGTCACTGACTGTGAGAG	1076
Q	y	963	TGTCACTTCCAAAGAGATCCGCGGCTGTGAGCTGCGGCTGGGGCTCCCGCTTCACGA	1022
D	b	1077	TGTCACTTCCAAAGAGATCCGCGGCTGTGAGCTGCGGCTGGGGCTCCCGCTTCACGA	1136
Q	y	1023	GTACCGTGATCTTATGTGACACCAATGCTGCTGCTGTGTGTGACACGCGGGACCGAGCTC	1082
D	b	1137	GTACCGTGATCTTATGTGACACCAATGCTGCTGCTGTGTGTGACACGCGGGACCGAGCTC	1196
Q	y	1083	CCGCACTTCTCCCGACCTGTACCTGTGGGCTCAAGATGTGAACGACGACAACTGTGAGAGACT	1142
D	b	1197	CCGCACTTCTCCCGACCTGTACCTGTGGGCTCAAGATGTGAACGACGACAACTGTGAGAGACT	1256
Q	y	1143	GCAGAGAAACAGGGTCAACCCACATCTTGAACATGGCCGGGAGATTGACAACTTCAACC	1202
D	b	1257	GCAGAGAAACAGGGTCAACCCACATCTTGAACATGGCCGGGAGATTGACAACTTCAACC	1316
Q	y	1203	TGAGCGGCTCAACCTACCAACAAATGGGGCCGTGGGATGAGGAGTGTGGCCGACGTGTGGCC	1262
D	b	1317	TGAGCGGCTCAACCTACCAACAAATGGGGCCGTGGGATGAGGAGTGTGGCCGACGTGTGGCC	1376
Q	y	1263	GCACGTGAGAGAGACGCAACCGCTTCAATTAGAGCTGTCAAAAGACACAGGGCACCAAGTGTCT	1322
D	b	1377	GCACGTGAGAGAGACGCAACCGCTTCAATTAGAGCTGTCAAAAGACACAGGGCACCAAGTGTCT	1436
Q	y	1333	GGTCCACTGCAAGATGGGGCTCAGCCGCTCAGCGGCGACACAGTGTGGGCTATGCCATGAA	1382
D	b	1437	GGTCCACTGCAAGATGGGGCTCAGCCGCTCAGCGGCGACACAGTGTGGGCTATGCCATGAA	1496
Q	y	1383	GCAGTAGCAATGTGACAGCCTGTGAGAGAGGCGCTGGCGCACGTGTGACAGAGCTTCGGGCCATATGC	1442
D	b	1497	GCAGTAGCAATGTGACAGCCTGTGAGAGAGGCGCGCTGGCGCACGTGTGACAGAGCTTCGGGCCATATGC	1556
Q	y	1443	CGGCGCCAAACCTGGGCTCCGTGGCGGACGCGAGATCTACAGGGGCATCTGACGAGGCGGCGAG	1497
D	b	1557	CGGCGCCAAACCTGGGCTCCGTGGCGGACGCGAGATCTACAGGGGCATCTGACGAGGCGGCGAG	1616
Q	y	1498	-----	1497
D	b	1617	CCGGCAGAGCCATGTCTGGGAGGAGAAAGTGGGTGGGCTGCCACAGAGAGACCCAGC	1676
Q	y	1498	-----GCCAGAACTGAGGGTGGTGGGAGGA	1524
D	b	1677	CCCTGGAAGTCTCTACACATTTCCACACTCTTCCGCCAGAACTGAGGGGTGGGTGGGCGGGA	1736
Q	y	1525	GAAAGTGTGAGCATGTGGAAGAGAGCAGGAGCGCCCGGAAGAGAGGCGTGGGCCACAGGGCC	1584
D	b	1737	GAAAGTGTGAGCATGTGGAAGAGAGCAGGAGCGCCCGGAAGAGAGGCGTGGGCCACAGGGCC	1796
Q	y	1585	ACGTATTAACCTCCGAGGGGTCAATGAGATGCATCACTTTCTGTGAGACCTCTCTTGAGACT	1644
D	b	1797	ACGTATTAACCTCCGAGGGGTCAATGAGATGCATCACTTTCTGTGAGACCTCTCTTGAGACT	1856
Q	y	1645	GGAGAGCACTCAGAGACCAAGTGCATGCGCAGAGCTTCTCTTCCAGAGTCTTTCACA	1704
D	b	1857	GGAGAGCACTCAGAGACCAAGTGCATGCGCAGAGCTTCTCTTCCAGAGTCTTTCACA	1916
Q	y	1705	TGAGAGGCTCTGCAGCCCTTCCACAGCTGTGAAGAGCAAGAGAGGCGACAGGTGTGA	1764
D	b	1917	TGAGAGGCTCTGTGAGGCTTCTCCACAGCTGTGCAAGAGCAAGAGAGGCGACAGGTGTGA	1976

[illegible]

RESULT 4			
AX451343	AX451343	1980 bp	DNA
LOCUS	Sequence 1	from Patent WO0224740.	linear
DEFINITION	AX451343		
ACCESSION	AX451343.1		
VERSION	AX451343.1	GT:21698394	
KEYWORDS	.		

SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1
AUTHORS LucHE,R.M. and Wei,B.
TITLE Dsp-15 dual-specificity phosphatase
JOURNAL Patent: WO 0224740-A 1 28-MAR-2002;
Ceptyr, Inc. (US)
FEATURES
Source Location/Qualifiers
1..1980
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 409 a 616 c 640 g 315 t
ORIGIN
Query Match 65.6%; Score 1774; DB 6; Length 1980;
Best Local Similarity 95.1%; Pred. No. 0;
Matches 1882; Conservative 0; Mismatches 0; Indels 98; Gaps 1;
QY 94 ATGGCCCTGGCTCAGTACAGTGGCGCTGCGCCCGGGCAGCGGGCTCCAGCGCCGTGGGG 153
Db 1 ATGGCCCTGGCTCAGTACAGTGGCGCTGCGCCCGGGCAGCGGGCTCCAGCGCCGTGGGG 60
QY 154 CCTGGGACCAAGCGCGTCCAGCGAAGAGTGCATCCAGCAAGGCAAGCTTTGCGGTG 213
Db 61 CCTGGGACCAAGCGCGTCCAGCGAAGAGTGCATCCAGCAAGGCAAGCTTTGCGGTG 120
QY 214 CTCCGTGGGGCTGCTCTGGGACTGCGAGATGGAGGGGCAATGATGACAGAGGCC 273
Db 121 CTCCGTGGGGCTGCTCTGGGACTGCGAGATGGAGGGGCAATGATGACAGAGGCC 180
QY 274 AGTTGTAGCCAAAGAGAGAGGCGCCGAGTGAAGAGAGGAGCTCCAGGGGACAGAGACAC 333
Db 181 AGTTGTAGCCAAAGAGAGAGGCGCCGAGTGAAGAGAGGAGCTCCAGGGGACAGAGACAC 240
QY 334 TTCGGGCAAGGATCCCAAGAGTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 393
Db 241 TTCGGGCAAGGATCCCAAGAGTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
QY 394 ATGTGACAGCTGCTGAGGCGCGAGATGACATCCGCTGGGACGCCAGCTGGAGGACACC 453
Db 301 ATGTGACAGCTGCTGAGGCGCGAGATGACATCCGCTGGGACGCCAGCTGGAGGACACC 360
QY 454 CGGCTCCCGGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 513
Db 361 CGGCTCCCGGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
QY 514 CAGATGAGAGCGGCTCTCTGGGCGTGGATTTCCTGACAGAGCTCCCGCAGCTGCACC 573
Db 421 CAGATGAGAGCGGCTCTCTGGGCGTGGATTTCCTGACAGAGCTCCCGCAGCTGCACC 480
QY 574 CTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 633
Db 481 CTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
QY 634 TTCACGTCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 693
Db 541 TTCACGTCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
QY 694 TGGGCGACACTCCAGTATTGACCAAGCATGTGAGGAGCTCTAGGAGGCGGCTTGTGTA 753
Db 601 TGGGCGACACTCCAGTATTGACCAAGCATGTGAGGAGCTCTAGGAGGCGGCTTGTGTA 660
QY 754 CCGGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 813
Db 661 CCGGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
QY 814 AGCTGCTCAATGAGTGAAGGCTATGGCGACCTGAGGAGCTCTGCTGCTGCTGCTGCTGCTGCTG 873
Db 721 AGCTGCTCAATGAGTGAAGGCTATGGCGACCTGAGGAGCTCTGCTGCTGCTGCTGCTGCTGCTG 780
QY 874 GAGCCTGGCGGGTCTCTCAGAACAGAGAGAGATGAGAGGCGATCCGTGCTGAGCTGTGG 933

Db 781 GAGCCTGGCGGGTCTCTCAGAACAGAGAGAGATGAGAGGCGATCCGTGCTGAGCTGTGG 840
QY 934 AAAGTGTGATGTAGTACCTGAGAGCTGCTACTTCCAAAGAGATCCGCGAGCTCTG 993
Db 841 AAAGTGTGATGTAGTACCTGAGAGCTGCTACTTCCAAAGAGATCCGCGAGCTCTG 900
QY 994 GAGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 1053
Db 901 GAGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 960
QY 1054 CTGCTGTGGGACAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1113
Db 961 CTGCTGTGGGACAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020
QY 1114 GAGTGAACGACGCAACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1173
Db 1021 GAGTGAACGACGCAACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
QY 1174 ATGGCCCGGAGATGACAACTTACCTGAGGCTGCTGAGGCTGCTGAGGCTGCTGAGGCTGCTG 1233
Db 1081 ATGGCCCGGAGATGACAACTTACCTGAGGCTGCTGAGGCTGCTGAGGCTGCTGAGGCTGCTG 1140
QY 1234 TGGATGAGAGTCCGCGCAGCTGCTGCGCACTGGAAGGAGAGCAGCGCTTCAATTGAG 1293
Db 1141 TGGATGAGAGTCCGCGCAGCTGCTGCGCACTGGAAGGAGAGCAGCGCTTCAATTGAG 1200
QY 1294 GCTGCAAGACAGAGGAGCAGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1353
Db 1201 GCTGCAAGACAGAGGAGCAGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
QY 1354 GCGGCGACAGTGTGGGCTATGGCTGTAAGCAGTACGATGAGCTGAGAGGCGGCTG 1413
Db 1261 GCGGCGACAGTGTGGGCTATGGCTGTAAGCAGTACGATGAGCTGAGAGGCGGCTG 1320
QY 1414 CGCCACGTGACAGGAGCTCGGCGCATCGCGCCGCGCCCAACCCCTGGCTTCTGCGCGCAGCTG 1473
Db 1321 CGCCACGTGACAGGAGCTCGGCGCATCGCGCCGCGCCCAACCCCTGGCTTCTGCGCGCAGCTG 1380
QY 1474 CAGATCTACAGAGGATCTCTGAGC----- 1497
Db 1381 CAGATCTACAGAGGATCTCTGAGC----- 1440
QY 1498 ----- 1497
Db 1441 GGTGGGCTCTCCCGAGAGAGGAGCAGCCAGCCCTGGAAGTCTTACACCACTTCCAGCTCTT 1500
QY 1498 --GCCAAGACCTGAGGCTGCTGCGGAGAGAGAGAGTGTAGGCAATGGAAGAGAGCAGGCA 1555
Db 1501 CGGCAGAACTGAGGCTGCTGCGGAGAGAGAGAGTGTAGGCAATGGAAGAGAGCAGGCA 1560
QY 1556 GCGCCGAAAGAGAGCTGCGGCGCACGCGCAGCTATTAACCTCCGAGGGGCTCATGAGTCC 1615
Db 1561 GCGCCGAAAGAGAGCTGCGGCGCACGCGCAGCTATTAACCTCCGAGGGGCTCATGAGTCC 1620
QY 1616 ATCAGTCTTTCGAGGCTCCCTGAGCTGGAGGACCACTGAGAGCAGAGGACATGCA 1675
Db 1621 ATCAGTCTTTCGAGGCTCCCTGAGCTGGAGGACCACTGAGAGCAGAGGACATGCA 1680
QY 1676 GAGGCTTCTCTTCCACAGAGTCTTACATGAAGAGCTCTGACAGCCCTTCCACAGGCTT 1735
Db 1681 GAGGCTTCTCTTCCACAGAGTCTTACATGAAGAGCTCTGACAGCCCTTCCACAGGCTT 1740
QY 1736 GCAAGGACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1795
Db 1741 GCAAGGACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1800
QY 1796 CAGTCAGTGTATCCCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1855
Db 1801 CAGTCAGTGTATCCCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1860
QY 1856 GAGCAGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1915

Db	1861	GAGCAGAGGACGAGGGGCGAGGGGCGACGGGGCGACGGGAGAGGCCCTGCATTTCTCTACGCCCAAG	1920
Qy	1916	TTCCCGAAGGTGTGTGACACAGGCGCGCTGCATGACAGTGGAGAGAGGGCGGCGCTGA	1975
Db	1921	TTCCGGAAGGTGTGTGACACAGGCGCGCTGCATGACAGTGGAGAGAGGGCGGCGCTGA	1980
RESULT 5			
LOCUS	AX398774	2322 bp	DNA
DEFINITION	Sequence 10 from Patent WO0220732.	linear	PAT 27-MAY-2002
ACCESSION	AX398774		
VERSION	AX398774.1	GI:21261307	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1		
AUTHORS	Liu,J.R.		
TITLE	Regulation of human map kinase phosphatase-like enzyme		
JOURNAL	Patent: WO 0220732-A 10 14-MAR-2002;		
	Bayer Aktiengesellschaft (DE)		
FEATURES	Location/Qualifiers		
source	1. 2322		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
BASE COUNT	467 a	751 c	718 g
ORIGIN		386 t	
Query Match	62.8%;	Score 1699;	DB 6;
Best Local Similarity	90.9%;	Pred. No. 0;	
Matches 1987;	Conservative	0;	Mismatches 75; Indels 123; Gaps 12.
Qy	132	CGGCGCTCCAGCCCGCTGGGGCCCTGGGAGCAGGGCGGTCCAGCGAAGAGTCCGACTCCA	191
Db	147	CGCCCTACCCCTGTGGGGCTCTCTTCGGCAGACCTGTGTCCAGGAGAGTCCGACTCCA	206
Qy	192	GCGAAGCAGAGCTTTGCGGTCTCC--GTGGGGCTGTCTCTGGACTGCAGATGAGGGG	250
Db	207	GGCA---AGAGCTTTGGCGTGTCTCTGTGGGGCTGTCTCTGGGACTGCAGATGAGGGG	262
Qy	251	ACAATGATATGCACGACGAGGCCAGTTCTGAGCCAAACAGAAAGCCCCGAGTGAAGG	310
Db	263	ACAATGATATGCAGCAGAGGCCAGTTCTGAGCCAAACAGAAAGCCCCGAGTGAAGG	322
Qy	311	AGTCCAGGGG---ACGAGACAGATT---CGGGCAGAGATCCAGAGTCCCAGAAAGA	365
Db	323	AGTCCAGGGGGAGCCAGACAGATTTTCGTGTCAGAGATCCAGAGTCCCAGAAAGA	382
Qy	366	GGAGGACGAGGACAGCACCTGCACCTCATGTGTACAGCTTGAGGCCCGCAGAGATGACAT	425
Db	383	GGAGGACGAGAGGACACCACTCATGTGTAGAGGCGGTGTGAGGCCCGCAGAGATGACAT	442
Qy	426	CGCGCTTGCA--GCCAGCTGAGAGCACCCCGGCTCTCCCGGCTCCGCTAAGTGC-----	478
Db	443	CGAGCTTGGAAAGCCCAAGCTGTGAGGACCCCGGGGCTCCCGGAGATCCGATACCTTGCT	502
Qy	479	TGGTATTTCTACAGAGAG--GAGAAAGTCTGAG--CCAGATGATAGAGGTCTCTCTGGG	536
Db	503	TGGTATTTCTACAGAGAAAGGAGAGGTCTGAGCCCAAGATGAAACGGTCTCTCTGGG	562
Qy	537	-CGTGATTTCCTGCACACAGTCCCGCAGCTGCACCCAGTCCAGCGTGTGCTGGCCCTCT	595
Db	563	ACGTGATTTCCCTGTACAGAGTCCCGCAGCTGCACCCAGTCCAGCGTGTGCTGGCCCTCT	622
Qy	596	GGAAGACACCCAGGTGTACTTATGAGAGACGGGGCTTTCAGCGTGTGCTGGGCG	655
Db	623	GGAGTGACACCCAGGTGTACTTATGAGAGACGGGGCTTTCAGCGTGTGCTGGGCG	682
Qy	656	AAAGCGGATCTTCAAGCCCATCTCATCAGACCATGTGTGGCCACACTTCAGATATTGC	715
Db	683	AAAGCGGATCTTCAAGCCCATCTCATCAGACCATGTGTGGCCACACTTCAGATATTGC	742

OY	716	ACCAAGCATGTGAGCAGCTCTTAGCGAGCGCCTTGTACCGGGTGGAGTGGCTCCACT	775
Db	743	ACCAAGCATGTGAGGAGAGCTTAGGGAGGCGGCTTGTACCGGGTGGAGTGGCTCCACT	802
OY	776	GGGCGACCACTACCGAGGAGAGCTACATCCGAGACAGAGCTCAATGATGAGGACG	835
Db	803	GGGCGACCACTACCGAGGAGAGCTACATCCGAGAGAGCTGCTCAATGATGAGACG	862
OY	836	CTATGGCGGACCTGGAGTCTCTCGGCGCTCCAGCGCGGAGCTGGCGGCTCTCAGAAC	895
Db	863	CTATGGCGGACCTGGAGTCTCTCGGCGCTCCAGCGCGGAGCTGGCGGCTCTCAGAAC	922
OY	896	AGGAGCAGATGAGAGCAGCGCATCCGTCGTGACTGTGGAAATGTTGGATGTCAGTGAC	955
Db	923	AGGAGCAGATGAGAGCGGGCCATCCGTCGTGACTGTGGAAATGTTGGATGTCAGTGAC	982
OY	956	TGGAGATGTACCTTCCAAAGAGATCCGCGAGGCTCTGAGATGGCGCTGGGCGCTCCCC	1015
Db	983	TGGAGATGTGCGCTTCCAGAGATCCGCGAGGCTCTGAGATGGCGCTGGGCGCTCCCC	1042
OY	1016	TCACAGCATACCGTGACTTATGACAAACCATGCTGTGCTGGTGGCACAAGGGGAC	1075
Db	1043	TCACAGCATACCGTGACTTATGACAAACCATGCTGTGCTGGTGGCACAAGGGGAC	1102
OY	1076	GAGCCTCCCGCATTTTCCCGCACCTCTACCTGGGCTCAGAGTGSAAACAGCAACTGG	1135
Db	1103	GAGCCTCCCGCATTTTCCCGCACCTCTACCTGGGCTCAGAGTGSAAACAGCAACTGG	1162
OY	1136	AGGAGCTGACAGAGACAGGGTACCCACATCTTGAACATGGCCCGGGAGATTGCAACT	1195
Db	1163	AGGAGCTGACAGAGAACAGGGTACCCACATCTTG--ATGGCCCGGGAGATTGCAACT	1219
OY	1196	TCTACCCGAGCGGTTACCTACCTACCAATGTGGCGCTGGGATGAGAGATCGGCGCAGC	1255
Db	1220	TCTACCCGAGCGGTTACCTACCTACCAATGTGGCGCTTGGGATGAGAGATCGGCGCAGC	1279
OY	1256	TGTCGCGCACTGGAGAGAGACGACCGCTTCAATGAGGCTGCANAGACAGGGCACCC	1315
Db	1280	TGTCGCGCACTGGAGAGAGACGACCGCTTCAATGAGGCTGCANAGAGACAGGGCACCC	1339
OY	1316	ACGTGCTGTGTCCACTGCAAAATGGGGGTACGCGGTCAAGGGGCAAGAGTGGGCTATG	1375
Db	1340	ACGTGCTGTGTCCACTGCAAAATGGGGGTACGCGGTCAAGGGGCAAGAGTGGGCTATG	1399
OY	1376	CCATGAACCACTAGCAATGCAAGCTGGAGAGGCGCTTGGCCACAGCTGCAGAGACTCGGC	1435
Db	1400	CCATGAACCACTAGCAATGCAAGCTGGAGAGGCGCTTGGCCACAGCTGCAGAGACTCGGC	1459
OY	1436	CCATGGCCCGCCCAACCCGTGGCTTCTGTGGCGAGTGTGAATCTTACAGAGGCAATCTGA	1495
Db	1460	CCATGGCCCGCCCAACCCGTGGCTTCTGTGGCGAGTGTGAATCTTACAGAGGCAATCTGA	1519
OY	1496	CG-----CG-----	1497
Db	1520	CGGCGACCGCGCAAGCATGTCTGGAGAGCAAAATGGGTGGGGTCTCCCAAGAGAGC	1579
OY	1498	-----GCAGACTGAGGCTGTG	1517
Db	1580	ACCGAGCCCTGGAAGTCTTACACATCTCCACCTTCTCCAGAGAACTGAGGGTGGTG	1639
OY	1518	GGGAGAGAAAGTTGTAGGCAATGGAGAGAGCGGACCGCGGAAAGAAAGAGCTGGGC	1577
Db	1640	GGGAGAGAAAGTTGTAGGCAATGGAGAGAGCGGACCGCGGAAAGAAAGAGCTGGGC	1699
OY	1578	CACGGCCACGATATTAACCTCCGAGGGGTCAATGAGGTCAATCACTTCTTGGAGCCCTCT	1637
Db	1700	CACGGCCACGATATTAACCTCCGAGGGGTCAATGAGGTCAATCACTTCTTGGAGCCCTCT	1759
OY	1638	TGGAGCTGGAGAGCACTCAGAGACCACTGACATGCCAGAGGCTTCTTCTCCACAGAT	1697
Db	1760	TGGAGCTGGAGAGCACTCAGAGACCACTGACATGCCAGAGGCTTCTTCTCCACAGAT	1819

QY	1698	CTACATGAGAGAGCCTCGACGCCCTCCACACACTTGGCAAGAGCCAGAGGAGGCCAGC	1757
Db	1820	CTTCACATGAAAGAGCCTCGACGCCCTTCCACACACTTGGACAGGACCAAGGAGGCCAGC	1879
QY	1758	AGGTGGACAGGGGGGCGCTCAGCCTGCCCTGAAGATCCGGCAGTCAGTGGTTACCTCCAGG	1817
Db	1880	AGGTGGACAGGGGGGCGCTCAGCCTGCCCTGAAGATCCGGCAGTCAGTGGTTACCTCCAGG	1939
QY	1818	GCAGTGGCGGTGGCCCAACCGGACCCAGCGCCTTCCAGAGAGTGAAGACAGGGGCGAGGGG	1877
Db	1940	GCAGTGGCGGTGGCCCAACCGGACCCAGCGCCTTCCAGAGAGTGAAGACAGGGGCGAGGGG	1999
QY	1878	AGGGGCGAGGAGAGCCCTGCATTTTCTCTACGCGCCAGGTTCCGGAAGGTGGTGAACACAG	1937
Db	2000	AGGGGCGAGGAGAGCCCTGCATTTTCTCTACGCGCCAGGTTCCGGAAGGTGGTGAACACAG	2059
QY	1938	CCAGGCTGCATGACAGTGGAGAGAGAGGGGCGAGGCTGTAGCCTTACACATGCCCCAGCTC	1997
Db	2060	CCAGGCTGCATGACAGTGGAGAGAGGGGCG - GGCTGTAGCCTTACACATGCCCCAGCTC	2117
QY	1998	CCCTCACACTGGAAGAGATCCACAACACTCTTGGAGAAACACCCTCAGCTGTGTTCCGCA	2057
Db	2118	CCCTCACACTGGAAGAGATCCACAACACTCTTGGAGAAACACCCTCAGCTGTGTTCCGCA	2177
QY	2058	CACATTCCTCTCAGTCCGCGCCCATACCCTGTACTACAGCTCACTCTCCACCCCTGTCA	2117
Db	2178	CACATTCCTCTCAGTCCGCGCCCATACCCTGTACTACAGCTCACTCTCCACCCCTGTCA	2237
QY	2118	CTACGGCCTTCACCTCCACCCCTGTCACTAAGCTCACAAGCTCACTCTCAAGCCTTAACTCCA	2177
Db	2238	CTACGGCCTTCACCTCCACCCCTGTCACTAAGCTCACAAGCTCACTCTCAAGCCTTAACTCCA	2297
QY	2178	GGCCCATGTCTGCTGTGTCCAAAGGC 2202	
Db	2298	GGCCCATGTCTGCTGTGTCCAAAGGC 2222	

RESULT 6	
BC004176	
LOCUS	BC004176 1905 bp mRNA linear PRI 12-JUL-2001
DEFINITION	Homo sapiens, similar to hypothetical protein FLJ10928, clone MGC:2772 IMAGE:2958967, mRNA, complete cds.
ACCESSION	BC004176
VERSION	BC004176.1 GI:13278815
KEYWORDS	MGC.
SOURCE	Homo sapiens.
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 1905) Strausberg,R. Direct Submission Submitted (01-MAR-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REFERENCE	NIH-MGC Project URL: http://mgc.nci.nih.gov
AUTHORS	Contact: MGC help desk Email: cgapbs-remail.nhl.gov
TITLE	Tissue Procurement: DCTD/DTP
JOURNAL	cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
REMARK	DNA Sequencing by: Institute for Systems Biology http://www.systemsbiology.org Contact: amadan@systemsbiology.org Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia Greene, Mark Kettelman and Anuradha Madan
COMMENT	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov Series: IRAL Plate: 2 Row: m Column: 8. Location/Qualifiers 1..1905
FEATURES	
source	

Query Match	Best Local Similarity	Matches 189: Conservative	61.6%	Score 1665.2	DB 9	Length 1905
			94.4%	Mismatches 0	Indels 8	Gaps 2
QY 907	GAGCAGGCATCCGTCGTCAGCTGTGGAAGGTTGATGTAGTACGTGGAGCTGTC	64				
DB 6	GAGGAGGCATCCGTCGT-AGCTGTGGAAGCTTGGATGTACGTACCTGGAGAGTGTG	64				
QY 967	ACTTCCAAAGAGATCCGCCAGGCTGTGGAGCTGGCCCTCCCTCCAGCATAC	1026				
DB 65	ACTTCCAAAGAGATCCGCCAGGCTGTGGAGCTGGCCCTCCCTCCAGCATAC	124				
QY 1027	CGTGACTTCATGACAACCAACAGATGCTGCTGCTGGTGGCAACGGGAGACCGACCTCCGC	1086				
DB 125	CGTGACTTCATGACAACCAACAGATGCTGCTGCTGGTGGCAACGGGAGACCGACCTCCGC	184				
QY 1087	ATCTTCCGCCACCTCTACCTGTGGCTCAGAGTGAAGCAGCAAACTCGAGAGCTGCAG	1146				
DB 185	ATCTTCCGCCACCTCTACCTGTGGCTCAGAGTGAAGCAGCAAACTCGAGAGCTGCAG	244				
QY 1147	AGGAACAGGGTACCCACATCTTGAACATGGCCGGGAGATTGAACATTTCTACCTGAG	1206				
DB 245	AGGAACAGGGTACCCACATCTTGAACATGGCCGGGAGATTGAACATTTCTACCTGAG	304				
QY 1207	CGCTTACCTACCAAAATGTGGCCTCTGGGATGAGAGTGGGCCACGTGCGCGAC	1266				
DB 305	CGCTTACCTACCAAAATGTGGCCTCTGGGATGAGAGTGGGCCACGTGCGCGAC	364				
QY 1267	TGGAAAGAGAGCAGCAGCTTCAATTGAGGCTGTCAGAGCAGCAAGGACCCAGCTGCTGTC	1326				
DB 365	TGGAAAGAGAGCAGCAGCTTCAATTGAGGCTGTCAGAGCAGCAAGGACCCAGCTGCTGTC	424				
QY 1327	CAGTCAGAGATGGGGGTCAAGCGGTCAGAGGGCAGAGTGGTGGCTTATGCCATGAAGAG	1386				
DB 425	CAGTCAGAGATGGGGGTCAAGCGGTCAGAGGGCAGAGTGGTGGCTTATGCCATGAAGAG	484				
QY 1387	TACGATGAGAGCCTGGAGCAGGCTTGGGCCACGTGACAGAGAACTTCGGGCCATCCCGC	1446				
DB 485	TACGATGAGAGCCTGGAGCAGGCTTGGGCCACGTGACAGAGAACTTCGGGCCATCCCGC	544				
QY 1447	CCCAACCCCTGGCTTCTGTGGCCAGCTGCAAGATCTACAGAGGATCTCTACG-----	1497				
DB 545	CCCAACCCCTGGCTTCTGTGGCCAGCTGCAAGATCTACAGAGGATCTCTACG-----	604				
QY 1498	-----	1497				
DB 605	CAGAGCATGTGTGGAGAGAGAAATGGTGGGGTCTCTCCAGAGAGACACCCAGCCCT	664				
QY 1498	-----GCCAGAACTGAGAGTGTGGTGGAGAGAG	1528				
DB 665	GAAATCTCTACACCATCCACACTTTCGCGAGAACTGAGGGTGTGGGAGAGAGAAG	724				

QY	1529	GTTGTAGCATGGAAGAAGACAGCCAGGCACGCCCCCGAAAGAAAGACCTGGGCCACAGGCACAGT	1588
Db	725	GTTGTAGCATGGAAGAAGACAGCCAGGCACGCCCCCGAAAGAAAGACCTGGGCCACAGGCACAGT	784
QY	1589	ATAAACCTCCGAGGGGGTATGAGGTCACATAGTCTCTCGAGAGCCCTCTCTGGAGCTGAG	1648
Db	785	ATAAACCTCCGAGGGGGTATGAGGTCACATAGTCTCTCGAGAGCCCTCTCTGGAGCTGAG	844
QY	1649	AGCACCTGAGAGACAGTACATGCGCAGAGGTCTTCTTCCACGAGTCTTACATGAA	1708
Db	845	AGCACCTGAGAGACAGTACATGCGCAGAGGTCTTCTTCCACGAGTCTTACATGAA	904
QY	1709	GAGCCTCTGACACCTTCTCCCAAGAGTTCGAAGAACCAGAGGGAGGCACGAGGTGGACAG	1768
Db	905	GAGCCTCTGACACCTTCTCCCAAGAGTTCGAAGAACCAGAGGGAGGCACGAGGTGGACAG	964
QY	1769	GGGCGCTCAGCCTCGCCTGAAGTCCCGCACAGTGAAGTTACCCTCCAGGGCAGTGGCGTG	1828
Db	965	GGGCGCTCAGCCTCGCCTGAAAGTCCCGCACAGTGAAGTTACCCTCCAGGGCAGTGGCGTG	1024
QY	1829	GTGGCCCAACCGAACCAGGCCTTCCAGAGACAGAGCAGGGGCGAAGGGGCGAGGGGA	1888
Db	1025	GTGGCCCAACCGAACCAGGCCTTCCAGAGACAGAGCAGGGGCGAAGGGGCGAGGGGA	1084
QY	1889	GAGCCCTGATTTTCTCTAGCGCCAGATTCCGGAAAGTGSTGAGCAGGCACAGCTGCAAT	1948
Db	1085	GAGCCCTGATTTTCTCTAGCGCCAGATTCCGGAAAGTGSTGAGCAGGCACAGCTGCAAT	1144
QY	1949	GACAGTGAAGAGAGAGGCGCAGGCGCTGAGCCCTCACATGCCCCAGCTCCCTGACACTG	2008
Db	1145	GACAGTGAAGAGAGAGGCGCAGGCGCTGAGCCCTCACATGCCCCAGCTCCCTGACACTG	1204
QY	2009	AAGAGATCCACAACTCTTTGAGAAACACCTTCAGTCTGTTGCGGCACACATTTCTCT	2068
Db	1205	AAGAGATCCACAACTCTTTGAGAAACACCTTCAGTCTGTTGCGGCACACATTTCTCT	1264
QY	2069	CAGCTCCGCGCCATTAACCCGCACTACAGCCTCAACCTCCACCCCTGTCACATACAGGCCTCA	2128
Db	1265	CAGCTCCGCGCCATTAACCCGCACTACAGCCTCAACCTCCACCCCTGTCACATACAGGCCTCA	1324
QY	2129	CTTCCACACCCCTTCACTACAGCCTCAACCTCTTAAGTCAAGCCAGGCCATGTCT	2188
Db	1325	CTTCCACACCCCTTCACTACAGCCTCAACCTCTTAAGTCAAGCCAGGCCATGTCT	1384
QY	2189	GCCCTGTCAGAGGGCTCAAGACTTTCTAAGTGGATGTGTAAGAGGACTAAGSTACCTT	2248
Db	1385	GCCCTGTCAGAGGGCTCAAGACTTTCTAAGTGGATGTGTAAGAGGACTAAGSTACCTT	1444
QY	2249	TGGGGGCAACAGCACCCTACTATTCTCAACCTTGAAGCCTGGACACTACCTGTGGCA	2308
Db	1445	TGGGGGCAACAGCACCCTACTATTCTCAACCTTGAAGCCTGGACACTACCTGTGGCA	1504
QY	2309	CGGATGAAGAAACAGAGCTTCCCGTGAAGAAAGGTCAAGCCTCCACCCCGCCCTCC	2368
Db	1505	CGGATGAAGAAACAGAGCTTCCCGTGAAGAAAGGTCAAGCCTCCACCCCGCCCTCC	1564
QY	2369	CTGCACTCTCTCTCTCTCCAGTTCAATTCTGGAACACGACAGGCCAAGCAACCAATGG	2428
Db	1565	CTGCACTCTCTCTCTCTCCAGTTCAATTCTGGAACACGACAGGCCAAGCAACCAATGG	1624
QY	2429	CCCCCAAGAGGCAAGATCTCAAGGCCACAGCCGGGGGAGGGTGTGAAGGGCTGGCAGA	2488
Db	1625	CCCCCAAGAGGCAAGATCTCAAGGCCACAGCCGGGGGAGGGTGTGAAGGGCTGGCAGA	1684
QY	2489	TGCGTTCCCTCATCCACTCCACCGGTCACAGSTTTTGTGCTCTCTCCACAGACTCTGT	2548
Db	1685	TGCGTTCCCTCATCCACTCCACCGGTCACAGSTTTTGTGCTCTCTCTCCACAGACTCTGT	1744
QY	2549	GACACCAAGCCAGATCACAGGGCACACGAGCCAGAGATAGTCTTTTGTCTTTCTGG	2608
Db	1745	GACACCAAGCCAGATCACAGGGCACACGAGCCAGAGATAGTCTTTTGTCTTTCTGG	1804
QY	2609	CCTCTGGCTAGTCAAGTTTTTTATAGCCTTAACAGTATCTGGCTTGTACTGAGAAATATAA	2666

DB	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	REMARK	COMMENT
Db	1805	CCTGCGCTAGTCAGTCTTTTCATAGCCTTACAGTATCGCTTTGTTACTGAGAAATATAA	1864										
Oy	2669	CACATTTTCATATAAAAAAAAAAAAAAAAAAAAAA	2704										
Db	1865	CACATTTTCATATAATTTGGTTAAAAAAAAAAAAAAAAAAAAA	1900										
RESULT 7	BC004210												
LOCUS	BC004210												
DEFINITION	Homo sapiens, similar to hypothetical protein FLJ10928, clone												
ACCESSION	MG:4436 IMAGE:2958967, mRNA, complete cds.												
VERSION	BC004210												
KEYWORDS	BC004210.1 GI:13278902												
SOURCE	MG.												
ORGANISM	Homo sapiens.												
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;												
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.												
TITLE	1 (bases 1 to 1905)												
JOURNAL	Strausberg, R.												
REMARK	Direct Submission												
COMMENT	Submitted (01-MAR-2001) National Institutes of Health, Mammalian												
	Gene Collection (MGC), Cancer Genomics Office, National Cancer												
	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,												
	USA												
	NIH-MGC Project URL: http://mgc.nci.nih.gov												
	Contact: MGC help desk												
	Email: cgapbs-remail.nih.gov												
	Tissue Procurement: DCTD/DTF												
	cDNA Library Preparation: Rubin Laboratory												
	DNA Sequencing by: Institute for Systems Biology												
	http://www.systemsbio.org												
	contact: amadane@systemsbiology.org												
	Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia												
	Greene, Mark Ketteman and Anuradha Madan												
	Clone distribution: MGC clone distribution information can be found												
	through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov												
	Series: IRAL Place: 10 Row: 9 Column: 1.												
FEATURES	Location/Qualifiers												
source	1..1905												
	/organism="Homo sapiens"												
	/db_xref="taxon:9606"												
	/clone="MG:4436 IMAGE:2958967"												
	/tissue_type="kidney, renal cell adenocarcinoma"												
	/clone_1lb="NIH_MGC_14"												
	/lab_host="DH10B-R"												
	/note="Vector: pOTB7"												
	272..1171												
	/product="start=1												
	/protein_id="AAH04210.1"												
	/db_xref="GI:13278903"												
	/translation="MAREIDNFPYPERFTYHNVRLMDESSAQLDPHKETHRFTIAARA												
	OGHIVLHCKMVCVSRAATVLAAYAKOECSEQLARVLOLRIRAKRNPGLFOLOID												
	YOGHILASRSHVMDROKVGVSPEHNPAPESLTPPLPPEPBGGEKVVGMDSQA												
	APKEEPGRPRNIRNGWVSIISLPSELESTSPMPFESSHESHEPQPPEP												
	QIARTGGGVNRGROPALKRSQSVTVTLGSA												

QY 967 ACTTCCAAAGATCCGCCAGGCTCTGGAGCTGGGCTCCCTCCAGAGTAC 1026
| | | | |
Db 65 ACTTCCAAAGATCCGCCAGGCTCTGGAGCTGGGCTCCCTCCAGAGTAC 124
| | | | |
QY 1027 GGTACCTTCATGACACACAGATGCTGCTGTGGTGCACAGCGGAGCCGACCTCCGC 1086
| | | | |
Db 125 GGTACCTTCATGACACACAGATGCTGCTGTGGTGCACAGCGGAGCCGACCTCCGC 184
| | | | |
QY 1087 ATCTTCCCCACCTCTACCTGGGCTCAGAGTGAAGCGAGCAACCTGGAGAGTGCAG 1146
| | | | |
Db 185 ATCTTCCCCACCTCTACCTGGGCTCAGAGTGAAGCGAGCAACCTGGAGAGTGCAG 244
| | | | |
QY 1147 AGGAACAGGGGTACCCACATCTTGAACATGAGCCCGGAGATTGACAACTTACCTGAG 1206
| | | | |
Db 245 AGGAACAGGGGTACCCACATCTTGAACATGAGCCCGGAGATTGACAACTTACCTGAG 304
| | | | |
QY 1207 CGCTTACCTACCAACATGTGCGCTCTGGGATGAGAGTGGCCAGCTGCTGCCGAC 1266
| | | | |
Db 305 CGCTTACCTACCAACATGTGCGCTCTGGGATGAGAGTGGCCAGCTGCTGCCGAC 364
| | | | |
QY 1267 TGGAGAGGAGACGACCGCTTCAATGAGGCTGACAGACACAGGCGACCCAGCTGTGTC 1326
| | | | |
Db 365 TGGAGAGGAGACGACCGCTTCAATGAGGCTGACAGACACAGGCGACCCAGCTGTGTC 424
| | | | |
QY 1327 CACTGCAAGATGGGCGTACGCGCTCAGCGGACAGTGTGCGCTATGCTATGAGAGAG 1386
| | | | |
Db 425 CACTGCAAGATGGGCGTACGCGCTCAGCGGACAGTGTGCGCTATGCTATGAGAGAG 484
| | | | |
QY 1387 TAGCAATGACAGCTTGAGACAGGCGCTGCGGACAGTGCAGAGTCCCGGCGCCGC 1446
| | | | |
Db 485 TAGCAATGACAGCTTGAGACAGGCGCTGCGGACAGTGCAGAGTCCCGGCGCCGC 544
| | | | |
QY 1447 CCCAACCTGGCTTCTCTGCGGCGAGCTGAGATCTACCGGCGATCTCTGAGC----- 1497
| | | | |
Db 545 CCCAACCTGGCTTCTCTGCGGCGAGCTGAGATCTACCGGCGATCTCTGAGCAGCCGC 604
| | | | |
QY 1498 ----- 1497
| | | | |
Db 605 CAGAGCCATGTCCTGGAGACAAAGTGGGTCTCCCGACAGAGACACCCAGCCCT 664
| | | | |
QY 1498 -----GCCAGAACCTGAGGGGTGGGAGAGAG 1528
| | | | |
Db 665 GAACTCTCTACACCATTTCCACCTTCTCCGCCAGAACCTGAGGGGTGGGAGAGAG 724
| | | | |
QY 1529 GTTGTAGGATGAGAGAGAGAGGCGAGGCGGCGGAAAGAGGCGGCGGCGAGCT 1588
| | | | |
Db 725 GTTGTAGGATGAGAGAGAGGCGAGGCGGCGGAAAGAGGCGGCGGCGAGCT 784
| | | | |
QY 1589 ATAAACCTCCGAGGGGTCATGAGTCCATCAGTCTTCTGAGCCCTCTTGGAGCTGAG 1648
| | | | |
Db 785 ATAAACCTCCGAGGGGTCATGAGTCCATCAGTCTTCTGAGCCCTCTTGGAGCTGAG 844
| | | | |
QY 1649 AGCACCTGAGAGACAGTACATGCGAGAGTCTTCTCTCCACAGATCTTACATGAA 1708
| | | | |
Db 845 AGCACCTGAGAGACAGTACATGCGAGAGTCTTCTCTCCACAGATCTTACATGAA 904
| | | | |
QY 1709 GAGGCTGCGAGCGCTTCCACAGCTTGAAGAGCAAGGAGGCGAGGCTGAGAGAG 1768
| | | | |
Db 905 GAGGCTGCGAGCGCTTCCACAGCTTGAAGAGCAAGGAGGCGAGGCTGAGAGAG 964
| | | | |
QY 1769 GGGCTCAGGCTGCGTGAAGTCCCGCAGTCACTGTGTTTACCTTCCAGAGGAGTCCCTG 1828
| | | | |
Db 965 GGGCTCAGGCTGCGTGAAGTCCCGCAGTCACTGTGTTTACCTTCCAGAGGAGTCCCTG 1024
| | | | |
QY 1829 GTGGCCAAACCGGAGCGCTTCCAGAGAGAGAGAGGCGGAGGCGAGGCGAGGGA 1888
| | | | |
Db 1025 GTGGCCAAACCGGAGCGCTTCCAGAGAGAGAGAGGCGGAGGCGAGGCGAGGGA 1084
| | | | |
QY 1889 GAGGCGCTCATTTCTCTACGCGCCAGGTTCCGAGAGGTTGAGAGAGGCGAGGCTGAT 1948
| | | | |
Db 1085 GAGGCGCTCATTTCTCTACGCGCCAGGTTCCGAGAGGTTGAGAGAGGCGAGGCTGAT 1144
| | | | |
QY 1949 GACAGTGAAGAGAGGCGGCTGAGCCCTCAGACATGCTCCAGGCTCCCTGACACTG 2008
| | | | |

Db 1145 GACAGTGAAGAGAGGCGGAGGCGCTGAGCCCTCAGACATGCCCAGGCTCCCTGACACTG 1204
| | | | |
QY 2009 AAGAGATCCACACACTCTTGGAGAGAACACCTCAGCTGTGTGGCGACATTTCTCT 2068
| | | | |
Db 1205 AAGAGATCCACACACTCTTGGAGAGAACACCTCAGCTGTGTGGCGACATTTCTCT 1264
| | | | |
QY 2069 CAGCTCCGCCCATACCTCTGAGAGAACACCTCAGCTGTGTGGCGACATTTCTCT 2128
| | | | |
Db 1265 CAGCTCCGCCCATACCTCTGAGAGAACACCTCAGCTGTGTGGCGACATTTCTCT 1324
| | | | |
QY 2129 CCTCCACCCCTGTCTACTACAGCCTCAGCTCCTTACAGCCTTAAGTCCAGGCGCATGCT 2188
| | | | |
Db 1325 CCTCCACCCCTGTCTACTACAGCCTCAGCTCCTTACAGCCTTAAGTCCAGGCGCATGCT 1384
| | | | |
QY 2189 GCTGTCCAAAGGCTCAGACTTTTAACTGGAGTGTGTAGAGGAGTGAAGTACTT 2248
| | | | |
Db 1385 GCTGTCCAAAGGCTCAGACTTTTAACTGGAGTGTGTAGAGGAGTGAAGTACTT 1444
| | | | |
QY 2249 TGGGCGCAACAGACACCTAGTTTCAATCTCAACTAGCCCTGACACTCAGCTGTGCA 2308
| | | | |
Db 1445 TGGGCGCAACAGACACCTAGTTTCAATCTCAACTAGCCCTGACACTCAGCTGTGCA 1504
| | | | |
QY 2309 CGGAATGAAAACAGAGCTTCCCGTGCAAAAAGGCTCAGCGCTCCACCCCGCCCTCC 2368
| | | | |
Db 1505 CGGAATGAAAACAGAGCTTCCCGTGCAAAAAGGCTCAGCGCTCCACCCCGCCCTCC 1564
| | | | |
QY 2369 CTGCACCTCTCTCTCTCTCCAGTTTCAATCTGGAACAGCAGGCGCAGGCAACAGTGG 2428
| | | | |
Db 1565 CTGCACCTCTCTCTCTCTCCAGTTTCAATCTGGAACAGCAGGCGCAGGCAACAGTGG 1624
| | | | |
QY 2429 CCCCCAAAGGAGGAGAGATCTCGAGCGCCAGCGCGGAGGCGGAGGAGGCTGGAGCA 2488
| | | | |
Db 1625 CCCCCAAAGGAGGAGAGATCTCGAGCGCCAGCGCGGAGGAGGCTGGAGGAGGCTGGAGCA 1684
| | | | |
QY 2489 TCGCTTCCCTCATCCACCTCCACCGGTCAGATCTTTGCTGCTGCCAGACCTCTGT 2548
| | | | |
Db 1685 TCGCTTCCCTCATCCACCTCCACCGGTCAGAGTCTTTGCTGCTGCCAGACCTCTGT 1744
| | | | |
QY 2549 GACACACAGCCAGATCACAGGCGACAGGCGCAGAGATGCTCTTTTGTCTTTCTGG 2608
| | | | |
Db 1745 GACACACAGCCAGATCACAGGCGACAGGCGCAGAGATGCTCTTTTGTCTTTCTGG 1804
| | | | |
QY 2609 CCTGCGCTAGTCACTTTTATAGCCTTACAGTATCTGCGCTTGTACTGAAATATAA 2668
| | | | |
Db 1805 CCTGCGCTAGTCACTTTTATAGCCTTACAGTATCTGCGCTTGTACTGAAATATAA 1864
| | | | |
QY 2669 CACATTTTCATTAATAAAAAAAAAAAAAAAAAA 2704
| | | | |
Db 1865 CACATTTTCATTAATAAAAAAAAAAAAAAAAAA 1900
| | | | |
RESULT 8
AX451362
LOCUS AX451362 1416 bp DNA linear PAT 03-JUL-2002
DEFINITION Sequence 20 from Patent WO0224740.
ACCESSION AX451362
VERSION AX451362.1 GI:21698396
KEYWORDS
SOURCE
ORGANISM house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 Luche, R.M. and Wei, B.
TITLE Dsp-15 dual-specificity phosphatase
JOURNAL Patent: WO 0224740-A 20 28-MAR-2002;
Ceptlyr, Inc. (US)
FEATURES
source location/Qualifiers
1..1416
/organism="Mus musculus"
/db_xref="taxon:10090"
BASE COUNT 288 a 450 c 449 g 229 t

LGVLPLMSDPTQVYLDDGDFSVYSGQSRIFKFIQIOTMWTALQVLIHOACEALGSG
LVPGSALITWASHAHOERLNSQSLNEMTADLESIRLSPACGSSSEDEOMQAIR
AEIKVLVDVSDLESVTSKEIRLOMSKSSSTVSSTTRCCSMHSGDRASRIRP
HLVYGEEMMANLEIOLRNRYTHILNAREIDNYPFRFYTHNRLINDEBSAQLLPRH
KETHRFTEAARAOCTHVHCKMGVSRSAATVLAIAKQIECSLEQALRHVQELRPTA
RPNFGFLRQIYQIGILTARF"

BASE COUNT 288 a 449 c 449 g 230 t

ORIGIN

Query Match 51.5%; Score 1391.4; DB 9; Length 1416;
Best Local Similarity 99.8%; Pred. No. 5.3e-278;
Matches 1414; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 94 ATGGCCCTGTCACAGATGAGCGCTTCGCCCGGAGCGGGGCTCCACGGCCGTGGGG 153
DB 1 ATGGCCCTGTCACAGATGAGCGCTTCGCCCGGAGCGGGGCTCCACGGCCGTGGGG 60
QY 154 CCTGGGACACAGGCGGTCACAGGAGAGTGCATCCAGCGAAGGACAGAGCTTTCGGTG 213
DB 61 CCTGGGACACAGGCGGTCACAGGAGAGTGCATCCAGCGAAGGACAGAGCTTTCGGTG 120
QY 214 CTCGGTGGGGCTGTCTCTGGGACTGAGAGTGGAGGGACATGATGATGACAGAGGCC 273
DB 121 CTCGGTGGGGCTGTCTCTGGGACTGAGAGTGGAGGGACATGATGATGACAGAGGCC 180
QY 274 AGTTCGTAGCCACAGAGAGAGGCCCGAGTGAAGAGAGCTCCAGGGGACAGAGAC 333
DB 181 AGTTCGTAGCCACAGAGAGAGGCCCGAGTGAAGAGAGCTCCAGGGGACAGAGAC 240
QY 334 TTCGGGACAGAGATCCACAGATGCCAGAAAGACAGAGAGAGAGACCTGACCTC 393
DB 241 TTCGGGACAGAGATCCACAGATGCCAGAAAGACAGAGAGAGAGACCTGACCTC 300
QY 394 ATGGATGACGCTGCTGAGGCGCCAGAGATGACATCCGCTGGAGAGCCAGCTGGAGCACC 453
DB 301 ATGGATGACGCTGCTGAGGCGCCAGAGATGACATCCGCTGGAGAGCCAGCTGGAGCACC 360
QY 454 CGGCTCCCGGCTCCGCTACCTGCTGATTTCTACACGAGAGAGAGAGTCTGAGC 513
DB 361 CGGCTCCCGGCTCCGCTACCTGCTGATTTCTACACGAGAGAGAGAGTCTGAGC 420
QY 514 CAGATGAGAGCGTCTCTGGGGCTGATTTCCCTGACAGAGCTCCCGAGCTGAC 573
DB 421 CAGATGAGAGCGTCTCTGGGGCTGATTTCCCTGACAGAGCTCCCGAGCTGAC 480
QY 574 CTGGGCTGGGCTCTGGGCTCTGGAGTACACCCAGCTGATTTCTGATGAGAGAGCGGGC 633
DB 481 CTGGGCTGGGCTCTGGGCTCTGGAGTACACCCAGCTGATTTCTGATGAGAGAGCGGGC 540
QY 634 TTCAGCTGAGCTGTGAGGCAAAAGCGGATCTTCAAGCCCATCTCCATCCAGCATG 693
DB 541 TTCAGCTGAGCTGTGAGGCAAAAGCGGATCTTCAAGCCCATCTCCATCCAGCATG 600
QY 694 TGGGCAACACTCCAGGATTTGACCAAGCATGTGAGGAGAGCTCTAGGAGGCGCTTGA 753
DB 601 TGGGCAACACTCCAGGATTTGACCAAGCATGTGAGGAGAGCTCTAGGAGGCGCTTGA 660
QY 754 CCGGCTGGACAGTCCCTACCTGGGCGACCATACAGAGAGAGCTCAACTCCGAAGAG 813
DB 661 CCGGCTGGACAGTCCCTACCTGGGCGACCATACAGAGAGAGCTCAACTCCGAAGAG 720
QY 814 AGCTGCTCAATGATGAGAGGCTATGGCGACCTGAGATCTTCCGCGCTCCAGCGCC 873
DB 721 AGCTGCTCAATGATGAGAGGCTATGGCGACCTGAGATCTTCCGCGCTCCAGCGCC 780
QY 874 GAGCTGGCGGCTCTTCAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 933
DB 781 GAGCTGGCGGCTCTTCAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 934 AAGGCTGATGATGAG 993
DB 841 AAGGCTGATGATGAG 899

QY 994 GAGCTGCCCTTGGGGCTCCCGCTCCAGAGATACCTGATCTTCATCGACAGACAGATGCTG 1053
DB 900 GAGCTGCCCTTGGGGCTCCCGCTCCAGAGATACCTGATCTTCATCGACAGACAGATGCTG 959
QY 1054 CTGCTGTGGGACAGC -GGAGCCAGGCTCCCGCATCTTCCCGCACCTCTACCTGGGCTC 1112
DB 960 CTGCTGTGGGACAGCAGCGGGGACCGAGGCTCCCGCATCTTCCCGCACCTCTACCTGGGCTC 1019
QY 1113 AGAGTGAACCGACGAACCTTGAGAGAGCTGACAGAGAAACAGAGGCTACCCACATCTTGA 1172
DB 1020 AGAGTGAACCGACGAACCTTGAGAGAGCTGACAGAGAAACAGAGGCTACCCACATCTTGA 1079
QY 1173 CATGCGCCGGAGATTTGACAACTTCTACCTGAGGCTTCACTTACCTACCAATGTGCGCT 1232
DB 1080 CATGCGCCGGAGATTTGACAACTTCTACCTGAGGCTTCACTTACCTACCAATGTGCGCT 1139
QY 1233 CTGGATGAGAGTGGGCGCCAGCTGCTCCGACATGGAAGAGAGAGAGAGAGAGAG 1292
DB 1140 CTGGATGAGAGTGGGCGCCAGCTGCTCCGACATGGAAGAGAGAGAGAGAGAGAG 1199
QY 1293 GCGTCAGAGAGACAGAGGACACCGACAGTGTGTCACCTGCAAGATGAGGCGCTG 1352
DB 1200 GCGTCAGAGAGACAGAGGACACCGACAGTGTGTCACCTGCAAGATGAGGCGCTG 1259
QY 1353 AGCGGCACAGTGTGCGCTATGCTGATGAGAGAGATGAGAGAGAGAGAGAGAGAG 1412
DB 1260 AGCGGCACAGTGTGCGCTATGCTGATGAGAGAGATGAGAGAGAGAGAGAGAGAG 1319
QY 1413 GCGCCAGCTGACAGAGCTCCGCGCCATGCGCCGCAACCTGTGCTTCTGCGGACCT 1472
DB 1320 GCGCCAGCTGACAGAGCTCCGCGCCATGCGCCGCAACCTGTGCTTCTGCGGACCT 1379
QY 1473 GCAGATCTACCGAGGCGCATCTGAGGCGCAGAACCTGA 1509
DB 1380 GCAGATCTACCGAGGCGCATCTGAGGCGCAGAACCTGA 1416

RESULT 10
AX398768 1755 bp DNA linear PAT 27-MAY-2002
LOCUS AX398768
DEFINITION Sequence 4 from Patent W00220732.
ACCESSION AX398768
VERSION AX398768.1 GI:21261303
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1
AUTHORS Liou,J.R.
TITLE Regulation of human map kinase phosphatase-like enzyme
JOURNAL Patent: WO 0220732-A 4 14-MAR-2002;
Bayer Aktiengesellschaft (DE)
FEATURES
source 1..1755
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 350 a 577 c 505 g 323 t
ORIGIN

Query Match 49.4%; Score 1337; DB 6; Length 1755;
Best Local Similarity 94.7%; Pred. No. 1e-266;
Matches 1427; Conservative 0; Mismatches 0; Indels 80; Gaps 1;

QY 1174 ATGGCCGGGAGATTGACAACTTACCTGAGCGCTTACCTACCAATGTGGGCTC 1233
DB 322 ATGGCCGGGAGATTGACAACTTACCTGAGCGCTTACCTACCAATGTGGGCTC 381
QY 1234 TGGGATGAGAGTGGGCGCAGTGTGCGGCACATGGAAGAGAGAGAGAGAGAGAG 1293
DB 382 TGGGATGAGAGTGGGCGCAGTGTGCGGCACATGGAAGAGAGAGAGAGAGAGAG 441
QY 1294 GCTGCAAGAGCACAGGAGCACCCAGCTGCTGCTCACTGCAAGATGGGCGTCA 1353

```

|||||
Db 442 GCTGCAAGAGCAGCAGGAGACCCACCTGCTGTGTCACATGCAAGATGGGCGTCAAGCCGCTCA 501
OY 1354 GGGGCGACAGTGTGCGCTATGCGATGAAGCAGTACGATGAGCCTGGAGAGGCGCTG 1413
Db 502 GGGGCGACAGTGTGCGCTATGCGATGAAGCAGTACGATGAGCCTGGAGAGGCGCTG 561
OY 1414 GGGCAGTGTGACAGAGCTCGCGGCCATCGCCCGCCCAACCCCTGTGCTTCTGCGCCAGCTG 1473
Db 562 GCGCCAGTGTGACAGAGCTCGCGGCCATCGCCCGCCCAACCCCTGTGCTTCTGCGCCAGCTG 621
OY 1474 CAGATCTACAGAGGATCTGTACAGGCGCAGAACCTGAGGTTGGTGGGAGAGAGAGTTGT 1533
Db 622 CAGATCTACAGAGGATCTGTACAGGCGCAGAACCTGAGGTTGGTGGGAGAGAGAGTTGT 681
OY 1534 AGGCATGGAAGAGAGCAGCAGGAGCCCGGAAAGAGAGAGCTGGGCGCAGGCGCATTA 1593
Db 682 AGGCATGGAAGAGAGCAGCAGGAGCCCGGAAAGAGAGAGCTGGGCGCAGGCGCATTA 741
OY 1594 CCTCGAGAGGATCATGAGTGTGATGATGCTTGTGAGAGCCCTCTTGTGAGCTGGAGAGC 1653
Db 742 CCTCGAGAGGATCATGAGTGTGATGATGCTTGTGAGAGCCCTCTTGTGAGCTGGAGAGC 801
OY 1654 CTCAGAGAGCAGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1713
Db 802 CTCAGAGAGCAGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 861
OY 1714 TCTGAGAGCCCTTCCACAGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1773
Db 862 TCTGAGAGCCCTTCCACAGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 921
OY 1774 TCAGCCTGCTGTAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1833
Db 922 TCAGCCTGCTGTAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 981
OY 1834 CAACCGGAGCAGGCTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1893
Db 982 CAACCGGAGCAGGCTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1041
OY 1894 CTGATTTTCTTCTACAGCCTGCGAAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAG 1953
Db 1042 CTGATTTTCTTCTACAGCCTGCGAAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAG 1101
OY 1954 TGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2013
Db 1102 TGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1161
OY 2014 GATCCACATCTCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2073
Db 1162 GATCCACATCTCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1221
OY 2074 CGGCGCCATACCGCTGCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2133
Db 1222 CGGCGCCATACCGCTGCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1281
OY 2134 CACCCCTGTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2193
Db 1282 CACCCCTGTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1341
OY 2194 TCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2253
Db 1342 TCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1401
OY 2254 GCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2313
Db 1402 GCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1434
OY 2314 TGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2373
Db 1435 -----CCTGCA 1441
OY 2374 CCTCTGTCTCTCTCCAGATTCATTCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2433

```

```

Db 1442 CCTCTGTCTCTCTCCAGTTCATTCCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1501
OY 2434 AAAGCAGAGCAGAGATCTCAGGCGCCAGCCAGCGGAGAGAGAGAGAGAGAGAGAGAGAG 2493
Db 1502 AAAGCAGAGCAGAGATCTCAGGCGCCAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1561
OY 2494 TCCCTCATCTACCTCAGCGGCTTCAGAGTGTGCTGTCTGTCTGTCTGTCTGTCTGTCT 2553
Db 1562 TCCCTCATCTACCTCAGCGGCTTCAGAGTGTGCTGTCTGTCTGTCTGTCTGTCTGTCT 1621
OY 2554 CACGCGAGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2613
Db 1622 CACGCGAGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1681
OY 2614 GCGTAGTCAGTTTTCATACCTTACAGTATCTGCTTTGTACTAGAGAAATTAACACAT 2673
Db 1682 GCGTAGTCAGTTTTCATACCTTACAGTATCTGCTTTGTACTAGAGAAATTAACACAT 1741
OY 2674 TTTTCATA 2680
Db 1742 TTTTCATA 1748

RESULT 11
AK001790
LOCUS
DEFINITION
Homo sapiens cDNA FLJ10928 fis, clone OVARC100473, weakly similar
to DUAL SPECIFICITY PROTEIN PHOSPHATASE 3 (EC 3.1.3.48) (EC
3.1.3.16).
ACCESSION
AK001790
VERSION
AK001790.1 GI:7023282
KEYWORDS
Oligo cloning; fls (full insert sequence).
SOURCE
Homo sapiens ovarian cancer cDNA to mRNA, clone_1lb:OVARC1
clone:OVARC100473.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1
Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
Nishikawa,T., Nagai,K., Sugano,S., Ishibashi,T., Fujimori,K.,
Tanai,H., Kimata,M., Matanabe,M., Hiraoka,S., Ishi,S., Kawai,Y.,
Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagahari,K.,
Masuko,Y. and Kanehori,K.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 1755)
Isogai,T. and Otsuki,T.
Direct Submission
Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute,
Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan: cDNA full insert
sequencing: Research Association for Biotechnology: cDNA library
construction, 5'- & 3'-end one pass sequencing and clone selection:
Helix Research Institute (supported by Japan Key Technology Center
etc.) and Department of Virology, Institute of Medical Science,
University of Tokyo.
FEATURES
source
1..1755
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="OVARC100473"
/issue_type="ovarian cancer"
/clone_1lb="OVARC1"
/note="cloning vector: pME18FLJ3"
232..657
/note="unnamed protein product"
/codon_start=1
/protein_id="BAA91913.1"
/db_xref="GI:7023283"
/translation="MEGTMMQORPVLSSOHPFSLTNSPPASHPAREIDNFPERFT
YHNVRIMDEBSQQLPHKETHRFLEARAGSTHVLVKNKGVSSAATVLAAMKQY

```

BASE COUNT 350 a 577 c 505 g 323 t
 ORIGIN
 Query Match 49.4%; Score 1337; DB 9; Length 1755;
 Best Local Similarity 94.7%; Pred. No. 1e-266;
 Matches 1427; Conservative 0; Mismatches 0; Indels 80; Gaps 1;

1174 ATGGCCCGGAGATTGACAACTTCTACCCCTGAGCGCTTCCATTACACACAAATGGCGCTC 1233
 322 ATGGCCCGGAGATTGACAACTTCTACCCCTGAGCGCTTCCATTACACACAAATGGCGCTC 381
 1234 TGGAGTAGAGAGTCCGCCAGCTGCTGCCGCACTGGAAGAGAGCAGCCGCTTCAATTGAG 1293
 382 TGGAGTAGAGAGTCCGCCAGCTGCTGCCGCACTGGAAGAGAGCAGCCGCTTCAATTGAG 441
 1294 GCTGGAAGAGCAGAGGCAACCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1353
 442 GCTGGAAGAGCAGAGGCAACCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 501
 1354 GCGCCACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1413
 502 GCGCCACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 561
 1414 GCGCCACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1473
 562 GCGCCACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 621
 1474 CAGATCTACAGAGGATCTGAGGCGGCAAGAACTGAGGCTGCTGCTGCTGCTGCTGCTG 1533
 622 CAGATCTACAGAGGATCTGAGGCGGCAAGAACTGAGGCTGCTGCTGCTGCTGCTGCTG 681
 1534 AGGCAATGGAAGAGAGCCAGGAGCGCCGAAAGAGAGGCTGCTGCTGCTGCTGCTGCTG 1593
 682 AGGCAATGGAAGAGAGCCAGGAGCGCCGAAAGAGAGGCTGCTGCTGCTGCTGCTGCTG 741
 1594 CCTCCGAGGGGTCATGAGGTCATGAGTCTTCTGAGAGCCCTCTTGGAGCTGGAGAGCAC 1653
 742 CCTCCGAGGGGTCATGAGGTCATGAGTCTTCTGAGAGCCCTCTTGGAGCTGGAGAGCAC 801
 1654 CTGACAGACCAATGATGATGAGGTCATGAGTCTTCTGAGAGCCCTCTTGGAGCTGGAGAG 1713
 802 CTGACAGACCAATGATGATGAGGTCATGAGTCTTCTGAGAGCCCTCTTGGAGCTGGAGAG 861
 1714 TCTGACAGCCCTTCCCAAGAGCTTGAAGAGCAAGGAGAGCCAGAGTGGAGAGAGGAGGCC 1773
 862 TCTGACAGCCCTTCCCAAGAGCTTGAAGAGCAAGGAGAGCCAGAGTGGAGAGAGGAGGCC 921
 1774 TTAGCCTGCTGAGAGTCCCGCAAGTCACTGCTTCCAGGAGAGTGGAGAGAGGAGGCC 1833
 922 TTAGCCTGCTGAGAGTCCCGCAAGTCACTGCTTCCAGGAGAGTGGAGAGAGGAGGCC 981
 1834 CAACCGAGAGCCAGGCTTCCAGAGAGAGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGCC 1893
 982 CAACCGAGAGCCAGGCTTCCAGAGAGAGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGCC 1041
 1894 CTGATTTCTCTTACAGCCAGGCTTCCAGAGAGTGGAGAGAGGAGGAGGAGGAGGAGGCC 1953
 1042 CTGATTTCTCTTACAGCCAGGCTTCCAGAGAGTGGAGAGAGGAGGAGGAGGAGGAGGCC 1101
 1954 TGGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2013
 1102 TGGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1161
 2014 GATTCACAACCTCTTGGAGAGAGCCCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2073
 1162 GATTCACAACCTCTTGGAGAGAGCCCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1221
 2074 CGGCCCCATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2133
 1222 CGGCCCCATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1281
 2134 CACCCCTGTCTACTAGAGCTCTCTCTTACAGCCTTAAGTCCAGAGGCCCATGTCTGCTG 2193

1282 CACCCCTGTCTACTACAGCCTCCTACAGCCTTAAGTCCAGAGGCCATGTGCTGCTG 1341
 2194 TCCAGAGGCTCAAGACTTCTTAAGTGGATGTGTGAGAGGAGTGAAGTACTTTGGGG 2253
 1342 TCCAGAGGCTCAAGACTTCTTAAGTGGATGTGTGAGAGGAGTGAAGTACTTTGGGG 1401
 2254 GCAACAGCACCCTAGTTTCTTAAGTGGATGTGTGAGAGGAGTGAAGTACTTTGGGG 2313
 1402 GCAACAGCACCCTAGTTTCTTAAGTGGATGTGTGAGAGGAGTGAAGTACTTTGGGG 1434
 2314 TGAAGAGAGGCTTCCGTTGCAAGAGGCTGCAAGGCTCCACCCCGCCCTCCCTGCA 2373
 1435 -----CCCTGCA 1441
 2374 CTTCTGCTCTCTCCAGTTCATTTCTGGAACAGCAGGCGCAGGCAACAGTGGCCCC 2433
 1442 CTTCTGCTCTCTCCAGTTCATTTCTGGAACAGCAGGCGCAGGCAACAGTGGCCCC 1501
 2434 AAAGCAGGAGAGATCTTCAGGCGCCAGGCGGAGAGGCTGGAAGGCTGGCAGATGCT 2493
 1502 AAAGCAGGAGAGATCTTCAGGCGCCAGGCGGAGAGGCTGGAAGGCTGGCAGATGCT 1561
 2494 TCCCTCATCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 2553
 1562 TCCCTCATCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 1621
 2554 CAGGCGAGATCAAGGAGGCGCAGGCGGAGAGATGCTTTTCTTCTTCTTCTTCTTCT 2613
 1622 CAGGCGAGATCAAGGAGGCGCAGGCGGAGAGATGCTTTTCTTCTTCTTCTTCTTCT 1681
 2614 GCTATGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2673
 1682 GCTATGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1741
 2674 TTTCTATA 2680
 1742 TTTCTATA 1748

RESULT 12
 AK000522
 LOCUS DEFINITION 2064 bp mRNA linear PRI 22-FEB-2000
 AK000522
 ACCESSION Homo sapiens cDNA FLJ20515 fls, clone KAT09889.
 AK000522.1 GI:7020673
 VERSION Homo sapiens signet-ring cell carcinoma cell_line:KATO III cDNA to
 KEYWORDS mRNA, clone_11b:KAT clone:KAT09889.
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (sites)
 Watanabe,K., Kumagai,A., Itakura,S., Yamazaki,M., Tashiro,H.,
 Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
 Nakamura,Y., Isogai,T. and Sugano,S.
 NEDO human cDNA sequencing project
 Unpublished
 2 (bases 1 to 2064)
 Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
 Shibahara,T., Tanaka,T. and Nakamura,Y.
 Direct Submission
 Submitted (15-FEB-2000) Sumio Sugano, Institute of Medical Science,
 University of Tokyo, Department of Virology, Shirokane-dai, 4-6-1,
 Minato-ku, Tokyo 108-8639, Japan (E-mail:cdna@ims.u-tokyo.ac.jp,
 Tel:81-3-5449-5286, Fax:81-3-5449-5416)
 NEDO human cDNA sequencing project supported by Ministry of
 International Trade and Industry of Japan; cDNA full insert
 sequencing: Research Association for Biotechnology; cDNA library
 construction, 5'- & 3'-end one pass sequencing: Department of
 Virology and Human Genome Center, Institute of Medical Science,
 University of Tokyo (partly supported by Science and Technology
 Agency).

FEATURES
source
Location/Qualifiers
1..2064
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="KAT09889"
/cell_line="KATO III"
/cell_type="signet-ring cell carcinoma"
/clone_lib="KAT"
/note="cloning vector pME18SFL3"
147..1331
/note="unnamed protein product"
/codon_start=1
/protein_id="BAA9128.1"
/db_xref="gi:7020674"
/translation="MALVTSRSPGSGASNPVGPMDQAVORSRRLORRSPFVLRGA
VLGDDGDNDNAARASSEPEKAPSEELHGDPTDPPGSGSPKOEOROHILMV
QLRQDDIRLALQLEAPRPRLRLYLIVSTRREGSLSDDEYLVLGVFPDSSPSC
LGLVPLMSDVOYLDGDSFVTSQSRLEKPLISIDTMATLVLVHQAACALGSG
LVPGSALTMAVSHYOERLNSQSEINWTAMADLSLRPSAEPPGSSSEOEOMEQAIK
AELMKYLESTSESDMEVFSHSSHEEPLOFPOLARTKGGQOYDRGPOPLKS
ROSVTLGSAVAVNRTOAFQEQGGGGGGEPCISSPFRKRVVROASVHDSGEEG
EA"

CDS
BASE COUNT 421 a 671 c 605 g 367 t
ORIGIN
Query Match 47.5%; Score 1285.6; DB 9; Length 2064;
Best Local Similarity 73.9%; Pred. No. 4.5e-256;
Matches 1998; Conservative 0; Mismatches 9; Indels 697; Gaps 1;

QY 1 CGTCTTCTGTCCTGCTGCGGGGTCCAGAGCTGTCCGCGGGTTAGAGGAGGGCCCTGCC 60
DB 54 CGTCTTCTGTCCTGCTGCGGGGTCCAGAGCTGTCCGCGGGTTAGAGGAGGGCCCTGCC 113
QY 61 CGGTGCGAGCCGAGTGTGCTGCGGCGCTGCTCATGGCCCTGTGTCACAGAGCCGTTGC 120
DB 114 CGGTGCGAGCCGAGTGTGCTGCGGCGCTGCTCATGGCCCTGTGTCACAGAGCCGTTGC 173
QY 121 CCCCCGGGACGCGCGCCCTCCAGCCCGTGGGGCCCTGGGACCAAGCGGTCCAGCGAAG 180
DB 174 CCCCCGGGACGCGCGCCCTCCAGCCCGTGGGGCCCTGGGACCAAGCGGTCCAGCGAAG 233
QY 181 AGTCGACTCCAGAGGAGGAGGAGCTTTGCGGTGCTCCGTGGGGCTGTCCCTGGGACTG 240
DB 234 AGTCGACTCCAGAGGAGGAGGAGCTTTGCGGTGCTCCGTGGGGCTGTCCCTGGGACTG 293
QY 241 GATGAGGAGGACATGATGATGACAGAGGAGCTTCTAGGCCAACAAGAGAGGCCCG 300
DB 294 GATGAGGAGGACATGATGATGACAGAGGAGCTTCTAGGCCAACAAGAGAGGCCCG 353
QY 301 AGTGAGGAGGAGCTCCAGGGGGACACAGAGACTTGGGGCAAGGATCCAGAGTCCCCAG 360
DB 354 AGTGAGGAGGAGCTCCAGGGGGACACAGAGACTTGGGGCAAGGATCCAGAGTCCCCAG 413
QY 361 AAGCAGAGAGAGAGAGGAGGAGGAGCTGACCTATGATGATGAGCTGAGAGCCGAGAGT 420
DB 414 AAGCAGAGAGAGAGAGGAGGAGGAGCTGACCTATGATGATGAGCTGAGAGCCGAGAGT 473
QY 421 GACATCGGCTGGGAGCGGAGCTGAGGAGCAACCGGCGCTCCCGGCTCCGCTACTGCTG 480
DB 474 GACATCGGCTGGGAGCGGAGCTGAGGAGCAACCGGCGCTCCCGGCTCCGCTACTGCTG 533
QY 481 GTAGTTCTACACAGAGAGAGAGGTCGTGAGCCAGATGAGAGGTCCTCTGCGCGTG 540
DB 534 GTAGTTCTACACAGAGAGAGAGGTCGTGAGCCAGATGAGAGGTCCTCTGCGCGTG 593
QY 541 GATTTCCTGACAGAGCTCCCGAGCTGACACCTGGGCGCTGGTCTTGCCCTTGAGAGT 600
DB 594 GATTTCCTGACAGAGCTCCCGAGCTGACACCTGGGCGCTGGTCTTGCCCTTGAGAGT 653
QY 601 GACACCCAGAGTACTAGATGAGAGCGGGGCTTACAGCGTGAGCTGAGTGGGCAAGC 660
DB 654 GACACCCAGAGTACTAGATGAGAGCGGGGCTTACAGCGTGAGCTGAGTGGGCAAGC 713

QY 661 CGGATCTTCAAGCCCATCTCCATCCAGACATGTGGGGCCACACACGATATGACACCA 720
DB 714 CGGATCTTCAAGCCCATCTCCATCCAGACATGTGGGGCCACACACGATATGACACCA 773
QY 721 GCATGTGAGGAGCTCTGAGGAGCGGCGCTTGTACCGGGTGGCAGTGCCTCAGCTGGGCC 780
DB 774 GCATGTGAGGAGCTCTGAGGAGCGGCGCTTGTACCGGGTGGCAGTGCCTCAGCTGGGCC 833
QY 781 AGCCACTACAGAGAGAGCTGAATCCGACAGAGAGCTGCTCAATGAGTGAAGGCTATG 840
DB 834 AGCCACTACAGAGAGAGCTGAATCCGACAGAGAGCTGCTCAATGAGTGAAGGCTATG 893
QY 841 GCCGACCTGAGTCTCTGCGGCGCTCCACAGCGGAGCGCTGGCGGGTCTCCAGAACAGAG 900
DB 894 GCCGACCTGAGTCTCTGCGGCGCTCCACAGCGGAGCGCTGGCGGGTCTCCAGAACAGAG 953
QY 901 CGATGAGAGCAGCGGATCCGTGCTGAGCTGTGAAGAGTTTGATGTGATGACCTGTGAG 960
DB 954 CGATGAGAGCAGCGGATCCGTGCTGAGCTGTGAAGAGTTTGATGTGATGACCTGTGAG 992
QY 961 AGTGTACTTCCAAAGAGATCCGCGAGGCTGTGAGCTGTGGGCTGCCCTCCAG 1020
DB 993 ----- 992
QY 1021 CAGTACCTGACTTCATGACACAACAGATGCTGTGTGTGGCACAGGGAGCCGAGCC 1080
DB 993 ----- 992
QY 1081 TCCCGATCTTCCCGACCTTACTGAGCTCAGATGAGAGGAGCAACACCTGGAGAG 1140
DB 993 ----- 992
QY 1141 CTGACAGAGAGAGGTCACCCACATCTTGAACATGCCCCGGAGATTGACAACTTTAC 1200
DB 993 ----- 992
QY 1201 CCTGAGCGCTTCACTACCAACATGTGGCGCTGTGGGATGAGGAGTGGGCCAGCTGTG 1260
DB 993 ----- 992
QY 1261 CCGACTGGAAGAGAGCAGCAGCGCTTCATTGAGGCTGCAAGAGACAGGGACCCACGTG 1320
DB 993 ----- 992
QY 1321 CTGCTCACTGCAAGATGGGCTTCAAGCCGCTCAGCGGCCACAGTGTGCTATGCGATG 1380
DB 993 ----- 992
QY 1381 AAGCAGTACGAATGACAGCGCTGGAGGAGCGGCGCTGCGCAGGTGAGAGGCTCCGGCCATC 1440
DB 993 ----- 992
QY 1441 GCCGCCCAACCTGCTTCTGCGCGAGCTGACAGATCTACAGGGGATCTCTGACGGCC 1500
DB 993 ----- 992
QY 1501 AGAAGCTGAGGCTGTGGGAGAGAGAGTTGATGATGGAAGAGAGCAGGAGCCCC 1560
DB 993 ----- 992
QY 1561 GAAAGAGAGCTGGGCGACGGCCAGCTATTAACCTCCGAGGGGTATGAGTCCATCAG 1620
DB 993 ----- 992
QY 1621 TCTTTGAGCCCTCTGAGAGCTGAGAGACCTCAGAGACAGTGAATGACAGAGT 1680
DB 993 -----TTGGAGCTGGAGAGCAGCTCAGAGACCAAGTGAATGACAGAGT 1036
QY 1681 CTCTCTTCCACAGAGTCTTACATGAAGGCTCTGAGCCCTTCCACAGCTTGAAG 1740
DB 1037 CTCTCTTCCACAGAGTCTTACATGAAGGCTCTGAGCCCTTCCACAGCTTGAAG 1096
QY 1741 GACCAAGGAGGCGCAGAGTGGAGAGGGGGCTCAGGCTGCGCTGAAGTCCCGCCAGTC 1800

|||||
Db 594 GATTCCCTGCACAGCTCCCGCAGCTGCACCCCTGGGCTGTGCTCCCTCTGGAGT 653
QY 601 GACACCCAGGTACTTATGATGGAGAGGGGCTTACGCTGAGCTGTGGGCAAGC 660
Db 654 GACACCCAGGTACTTATGATGGAGAGGGGCTTACGCTGAGCTGTGGGCAAGC 713
QY 661 CGGATCTTCAAGCCCATCTCCATCCAGACATGTGGGCGACACTCCAGATTTGGACCA 720
Db 714 CGGATCTTCAAGCCCATCTCCATCCAGACATGTGGGCGACACTCCAGATTTGGACCA 773
QY 721 GCATGTGAGGAGCTCTAGGAGCGGCTTGTACCGGCTGGCAGTGGCTTACCTGAGG 780
Db 774 GCATGTGAGGAGCTCTAGGAGCGGCTTGTACCGGCTGGCAGTGGCTTACCTGAGG 833
QY 781 AGGCACTACGAGAGAGCTGAACCTCCAGACAGAGCTGCTCAATGATGAGGAGCTATG 840
Db 834 AGGCACTACGAGAGAGCTGAACCTCCAGACAGAGCTGCTCAATGATGAGGAGCTATG 893
QY 841 GCCGACCTGAGGTCTCTGCGGCTTCCAGCGCCGAGCCTGGCGGGTCTCTAGAAACAGAG 900
Db 894 GCCGACCTGAGGTCTCTGCGGCTTCCAGCGCCGAGCCTGGCGGGTCTCTAGAAACAGAG 953
QY 901 CAGATGAGAGCGGATCCGTGCTGAGCTGTGGAAAGTGTGGATGTAGTACCTGGAG 960
Db 954 CAGATGAGAGCGGATCCGTGCTGAGCTGTGGAAAGTGTGGATGTAGTACCTGGAG 992
QY 961 AGTGTCACTTCCAAAGAGATCCGCGAGGCTGTGAGCTGCGGCTGGGCTCCCTCCAG 1020
Db 993 ----- 992
QY 1021 CAGTACCTGACTTCAATGACAACCATGCTGTGTGGGACAGCGGAGCCAGAGCC 1080
Db 993 ----- 992
QY 1081 TCCCGCATTTCCCCACCTCTACTGCGGCTCAGAGTGGAAACGACGAACTGGAGGAG 1140
Db 993 ----- 992
QY 1141 CTGCAGAGAACAGGGTCAACCCACATCTTGAACATGGCCGAGATTGACAACTTCTAC 1200
Db 993 ----- 992
QY 1201 CCTGAGCCCTTACCTACCAACATGTGCGCTGTGGGATGAGGAGTCGAGCCAGCTGCTG 1260
Db 993 ----- 992
QY 1261 CCGCAGCTGGAAGGAGCAGCAGCGCTTCAATTGAGGCTGCAAGAGCAGAGGACCCAGCTG 1320
Db 993 ----- 992
QY 1321 CTGTGCACTGCANAGATGGGGTCAAGCCGCTCAGCGGCCACAGTGTGGCTATGCCATG 1380
Db 993 ----- 992
QY 1381 AAGCAGTACGAATGCAGCTGGAGCAGCGCCCTGCGCCACGTGCAGAGCTCCGCCCATC 1440
Db 993 ----- 992
QY 1441 GCCGCCCCCAACCTGGGCTTCTGCGCAGCTGAGATCTACAGAGGATCTCTGAGGCGC 1500
Db 993 ----- 992
QY 1501 AGAACCTGAGGGTGGTGGGAGAGAAAGTTGTAGCATGGAAGAGAGCCAGAGCCCC 1560
Db 993 ----- 992
QY 1561 GAAGAAGAGCCTGGGCAAGGCCACGTATAAACCCTCCGAGGGGTCAAGAGTCCATCAG 1620
Db 993 ----- 992
QY 1621 TCTTCTGAGGCCCTCTTGGAGCTGGAGAGCATCTCAGAGACCATGACATGCCAGAGT 1680
|||||

Db 993 -----TTGAGCTGGAGAGCACTCAGACACCACTGACATGCCAGAGT 1036
QY 1681 CTTCCTTCCACAGAGTCTTACATGAAGAGCTCTGAGCCCTTCCACAGCTTGCAG 1740
Db 1037 CTTCCTTCCACAGAGTCTTACATGAAGAGCTCTGAGCCCTTCCACAGCTTGCAG 1096
QY 1741 GACCAAGGGAGCCAGCAGGTGGAGAGGGGCTCAGCTTCCCTGAAGTCCCGCACTC 1800
Db 1097 GACCAAGGGAGCCAGCAGGTGGAGAGGGGCTCAGCTTCCCTGAAGTCCCGCACTC 1156
QY 1801 AGTGTACCTCCAGGGCAGTGGGCTGCGGCAACCGGACCCAGGCTTCCAGAGCA 1860
Db 1157 AGTGTACCTCCAGGGCAGTGGGCTGCGGCAACCGGACCCAGGCTTCCAGAGCA 1216
QY 1861 GGAGCAGGGGAGGGGCGAGGGGAGAGAGCCCTGCACTTCTCTACTAGCCAGGTTCCG 1920
Db 1217 GGAGCAGGGGAGGGGCGAGGGGAGAGAGCCCTGCACTTCTCTACTAGCCAGGTTCCG 1276
QY 1921 GAAAGGTGTAGAGACAGGCGCAGCTGATGACAGTGGAGAGAGGGGCGAGGCTTGAGCTT 1980
Db 1277 GAAAGGTGTAGAGACAGGCGCAGCTGATGACAGTGGAGAGAGGGGCGAGGCTTGAGCTT 1336
QY 1981 CACACATGCGCAGGCTCCCTGACACCTGAGAGAGATCCCAACTCTTGGAGAAACACCC 2040
Db 1337 CACACATGCGCAGGCTCCCTGACACCTGAGAGAGATCCCAACTCTTGGAGAAACACCC 1396
QY 2041 TCAGCTGTGGCGCACACATCTCTCAGCTCGGCCGCCCATACCCGCTCAGTACAGGCTC 2100
Db 1397 TCAGCTGTGGCGCACACATCTCTCAGCTCGGCCGCCCATACCCGCTCAGTACAGGCTC 1456
QY 2101 ACCTCCACCCCTGTCACTAGGCGCTCAGCTCCACCCCTGTCACTACAGGCTCAGCTCC 2160
Db 1457 ACCTCCACCCCTGTCACTAGGCGCTCAGCTCCACCCCTGTCACTACAGGCTCAGCTCC 1516
QY 2161 TACAGCTTAAAGTCCACAGGCCATGTGCTGTCCAGAGGCTCAAGACTTCTAACTGG 2220
Db 1517 TACAGCTTAAAGTCCACAGGCCATGTGCTGTCCAGAGGCTCAAGACTTCTAACTGG 1576
QY 2221 GATGTGTAGAGGAGCTGAAGTACCTTGGGGGACACAGCAGCCCTAGTTCAATTCAC 2280
Db 1577 GATGTGTAGAGGAGCTGAAGTACCTTGGGGGACACAGCAGCCCTAGTTCAATTCAC 1636
QY 2281 CTCTAGCCCTGCACACTACCTGTGGCAGCGAATGAACAGAGCTTCCGCGCAAAAG 2340
Db 1637 CTCTAGCCCTGCACACTACCTGTGGCAGCGAATGAACAGAGCTTCCGCGCAAAAG 1696
QY 2341 GGTACGCTCCACACCCCGCCCTCCCTGACACTCTGTCTCTCTCCAGTTCATTCT 2400
Db 1697 GGTACGCTCCACACCCCGCCCTCCCTGACACTCTGTCTCTCTCCAGTTCATTCT 1756
QY 2401 GGACACAGCCAGGCGCAGCAACCAAGTGGGCCCCCAAGGCGAGGAGATCTCAGGCCCA 2460
Db 1757 GGACACAGCCAGGCGCAGCAACCAAGTGGGCCCCCAAGGCGAGGAGATCTCAGGCCCA 1816
QY 2461 GCCGCGGAGAGGTGGAAGGCTGGGAGATCGTTCCCTCATCCACCGGTCCAGG 2520
Db 1817 GCCGCGGAGAGGTGGAAGGCTGGGAGATCGTTCCCTCATCCACCGGTCCAGG 1876
QY 2521 TCTTGTGCTGTGCTCCAGAGCTCTCTGACACCAAGCGCAGATCAGAGGAGCCAGGCA 2580
Db 1877 TCTTGTGCTGTGCTCCAGAGCTCTCTGACACCAAGCGCAGATCAGAGGAGCCAGGCA 1936
QY 2581 GAGATAGTCTTCTTCTTCTCTTCTGAGCTGTGGCTGTGATGATTTTTCATAGCTTACA 2640
Db 1937 GAGATAGTCTTCTTCTTCTTCTTCTGAGCTGTGGCTGTGATGATTTTTCATAGCTTACA 1996
QY 2641 GTATCTGGCTTTGTACTGAGAAATAAACACATTTTCAATTAATAAAAAAAAAAAAA 2700
Db 1997 GTATCTGGCTTTGTACTGAGAAATAAACACATTTTCAATTAATAAAAAAAAAAAAA 2056
QY 2701 AAAA 2704
Db 2057 AAAA 2060

RESULT 14
BC028922 2736 bp mRNA linear ROD 07-AUG-2002
DEFINITION Mus musculus, similar to slingshot 3, clone MGC:25738
IMAGE:3987714, mRNA, complete cds.
ACCESSION BC028922
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2736)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (01-MAY-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amgdbcm.tmc.edu
Guanarane, P.H., Garcia, A.M., Lu, X., Huylk, S.W., Hale, S.M., Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAX Plate: 30 Row: 9 Column: 9
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarity but not identity to protein.
Location/Qualifiers
1. 2736
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="CZECH II"
/clone="MGC:25738 IMAGE:3987714"
/tissue_type="Mammary tumor metastasized to lung. Tumor arose spontaneously from a senescent normal mammary (clonal) outgrowth infected with the virus MMTV."
/clone_lib="NCI CGAP_Lu29"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
93. 2042
/codon_start=1
/product="similar to slingshot 3"
/protein_id="AAH28922.1"
/db_xref="GI:20810382"
/translation="MALVTVSPPASGSHSTPVPDQVVRRRGLQRRSPAVLRGAVLGLDGGDSNVASDEADSEBPSEEDPTEDOTKGGDLQSPKQVQRHLLMYELLRPODDIRLAOLEAARPPRLRLVAVSTGELESEPAILLGDFPSSHSGTGLVLPIMSPDYVLDGGGFSVNSGGOSRFFKFSVSTOTMATTLOVHOACSVASGLVPGSALNATYVOEKLNSDOCLNEMMSMDSFRRPRAEPGASDEOKMOALIAELWOYLDITDLSVTSKRETRQALELRIGCTPLOOTYRPTIDNOMLLMAQDRASTIPPHLYLGSWMANLEELQKNRSHILNNAARELDNFPFRFTYNNRVWDEESAQLPFWKETHRTFEDAAQGTQRLVLCCKMGVSSAAVLAAMKQVMDLEQALIHQEDRPTVVRNHFPTROLRTYQGITLASRQSHWQKAVVPEEPKLAPEVSLPLPPEEGSGGEVNVMLGEGSEPTKEELGRPRINLRGVMRSILPESESPPEAGGLPEVPSDEPLHPFQSLBRAGCGQVRKRGPRPALKRSQSVLAHSAALVASTRAFPQEGGQGOEBSERGMSSTPRLRKVKWRQASVDDSDREEDKA"
BASE COUNT 599 a 780 c 768 g 589 t
ORIGIN

Query Match 42.7%; Score 1155.2; DB 10; Length 2736;
Best Local Similarity 69.2%; Pred. No. 4,8e-229;
Matches 1947; Conservative 0; Mismatches 633; Indels 232; Gaps 18;
5 CTTCTGTCCTCCGCGGTCAGAGACTGTCGCGGGGTGAGGAAGGGCGTCCCGGT 64
1 CTTCTGTCCTCCGCGGTCAGAGACTGTCGCGGGGTGAGGAAGGGCGTCCCGGT 60
65 GCCAGCCGAGGTCCTGCGGGCTGCTCC--ATGGCCCTGTCACAGTACGCTTGC 121
61 GTTCCGCGGTCCTGCGGGCTGCTCC--ATGGCCCTGTCACAGTACGCTTGC 120
122 CCGCGGTCCTGCGGGCTGCTCC--ATGGCCCTGTCACAGTACGCTTGC 178
121 CCGCGGTCCTGCGGGCTGCTCC--ATGGCCCTGTCACAGTACGCTTGC 180
179 GAGTCTGACTCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 238
181 GAGGCGGTCCTGCGGGCTGCTCC--ATGGCCCTGTCACAGTACGCTTGC 240
239 AGGATGAG 298
241 AGGATGAG 300
299 CGAGTGGAG 358
301 CAGGTGAG 360
359 AG 418
361 GGAAGC---AGTGGCAAAAGCGGACAGCTGACCTGAGTGGAGTACAGAGAG 417
419 ATGACATCGGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 478
418 ATGATATCGGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 477
479 TGGTATGTCCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 538
478 TGGTATGTCCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 528
539 TGGATTTCCCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 598
529 TGGATTTCCCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 588
599 GTGATATCGGCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 648
589 GTGATATCGGCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 648
659 GCGGATCTTCAAGCCCATCTCCATCCAGAGAGAGAGAGAGAGAGAGAGAGAG 718
649 GCGGATCTTCAAGCCCATCTCCATCCAGAGAGAGAGAGAGAGAGAGAGAGAG 708
719 AAGCATGTGAGGACAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 778
709 AAGCATGTGAGGACAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 768
779 CCAGGACATGAG 838
769 CTAATGATGAG 828
839 TGGCGGACATGAG 898
829 TGTCTGACCTGAG 888
899 AGCAGATGAG 958
889 AAGAGATGAG 948
959 AAGAGATGAG 1018
949 AAGAGATGAG 1008

QY 1019 AGCAGTACCGTGAATTCATCGAACAACAGATGCTGCTGCTGTCACAGCGGGACCGAG 1078
 Db 1009 AGCAGTACCGTGAATTCATCGAACAACAGATGCTGCTGCTGTCACAGCGGGACCGAG 1068
 QY 1079 CCTCCCGATCTTCCCGACCTCTACCTGGGCTGACAGTGGAGCAGCAAACTGGAGG 1138
 Db 1069 CCTCCCGATCTTCCCGACCTCTACCTGGGCTGACAGTGGAGTGGCAACTGGAGG 1128
 QY 1139 AGCTGAGAGAAACAGGCTGACCCACATCTTGACATGAGCGCCGGAGATTGCACTTCT 1198
 Db 1129 AACTTCAGAAAAACAGAGTAACTACATCTGACATGAGCGCCGGAGATTGCACTTCT 1188
 QY 1199 ACCGTGAGCGCTTCACTACCAACAATGTCGCTCTGGAGTGAAGATGCGCCAGCTGC 1258
 Db 1189 TCCCTGAGCGCTTCACTATTTCAATGTACGTCTGGATGAGAAATCGCACAGCTTTC 1248
 QY 1259 TGCCCATCTGGAAGAGACGACCGCTTCAATGAGCTGCAAGACACAGGGCACCAG 1318
 Db 1249 TGCCCATCTGGAAGAGACACATGATTCATGAGAGTGGCCACAGCAGCGGCACTGGG 1308
 QY 1319 TGCTGCTCCACTGCAAGATGGGCGTCAAGCGCTCAAGCGGCAAGTGTGGCTATGCCA 1378
 Db 1309 TGCTGCTCCACTGCAAGATGGGCGTCAAGCGCTCAAGCGGCAAGTGTGGCTATGCCA 1368
 QY 1379 TGAAGCATGATGCAATGCAAGCTGAGAGAGCGCTGCGCCACTGCAAGAGCTCGGCCCA 1438
 Db 1369 TGAAGCATGATGCTGGAGCTGGAGCAAGCCCTGATCCAGTGCAGAGGAGCTCGGCCCA 1428
 QY 1439 TGCCCGCGCCCAACCTGCTCTGCGGCACTGCAATACCAAGGCACTCTGACGG 1498
 Db 1429 TTGTAGCGCCCAACCAAGGCTTCTGCGGCACTGCAATACCAAGGCACTCTGACGG 1488
 QY 1499 ----- 1498
 Db 1489 CCAGCGGAGAGACCACTGCTGGAGACAGAAAGTGGGTGCTGCCACAGAGAGACCCC 1548
 QY 1499 -----CCAGACCTGAGGGGCGGGG 1520
 Db 1549 TGCGCCCTGAAGTTCTTACACCATTTGCCACCTCTCCACCAAGGAGGAGTGGGG 1608
 QY 1521 AGGAGAAAGTTGATGAGCATGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1580
 Db 1609 AGGTGATGATGATGAGGCTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1668
 QY 1581 GGGCAGGTATAACCTCGAGAGGGGTCATGAGTCCATGATCTTCTGAGAGCCCTCTTG 1640
 Db 1669 GGGCAGGTATAACCTCGAGAGGGGTCATGAGTCCATGATCTTCTGAGAGCCCTCTG 1725
 QY 1641 AGCTGAGAGACACTGAGAGACAGTGCATGCGAGAGGCTTCTCTCCAGAGAGCTT 1700
 Db 1726 AGTCAAGAGATCCCGAGAGGCGGAGGCTTCCAGAGAGTCTTCTCTCTCTCTCTCT 1773
 QY 1701 CACATGAAGAGCTTCTGAGAGGCTTCCAGAGGCTTCCAGAGAGAGAGAGAGAGAGAGAG 1760
 Db 1774 CAGATGAAGAGCTTCTTCAATCTCTCTCAAGGCTTCCAGAGAGAGAGAGAGAGAGAG 1833
 QY 1761 TGGACAGGGGCTCGAGGCTGCGCTGAAGTCCGCAAGTCAAGTGTACCTTCCAGGGCA 1820
 Db 1834 TCCGCAAGGCGCTTGGCGCTGCGCTGAAGTCTGAGAGTCTGTTGCTTCAATAGTG 1893
 QY 1821 GTCCCGTGTGGGCAACCGGAGCGCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1880
 Db 1894 CTCCCGTGTGGGCAACCGGAGCGCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1947
 QY 1881 GGCAGGAGAGAGCTGATTTCTTACGCGGAGTTCGAGAGAGAGAGAGAGAGAGAGAG 1940
 Db 1948 AGCAGAGTGAAGCTGTGATTTCTTCAAGCTTACGCTCGGAGAGAGATGAGAGAGAGCA 2007
 QY 1941 GCGTGCATGACATGAG 1999
 Db 2008 GTGTAGATGACAG 2067
 QY 2000 CTGACACTGAAGAGATCCAAACTCTTGGG- GAAACACCCCTCACGCTGTGTTGCCGAC 2058

Db 2068 ACATTAATGCGAGCTGTGTCAGAGCGCTTCAGATGAGTGTCTTACACTTATGGCGCTGC 2127
 QY 2059 ACATTCTCTGAGCTCCGCGCCCATACCGTCACTACAGCCCTACCTCCACCCCTGTAC 2118
 Db 2128 ACAGCCCTGAGAGCTCTTCCA----- 2150
 QY 2119 TAGGCGCTACCTCCACCCCTGTCTACTACAGCTTCACTCTCTACAGCTTAAGTCCAG 2178
 Db 2151 -----GCTGCCCTCAGACCTTGGCCTTCTGAGGCTTAAGTCTAG 2194
 QY 2179 GCCCATGTCTGCTGCTCCAGAGGCTCAAGACTTCTTAACGGAGATGTGAGAGAGAC-T 2237
 Db 2195 CCATATATCTACTGTTAAGGCTCAATCTTTTCTACATAGAGTCTGGTTAAGCGGT 2254
 QY 2238 GAAGTACCTTTGGGGCAACAGACACCTAGTTTCAATCTCAAGTCTAGCCCTGCACACT 2297
 Db 2255 TTAAAGTGTCTTAAAGGTATTAACACCTAGTTTCAATCTAGTACACTCACTTAACATT 2314
 QY 2298 CACTGTGACAGCAATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2357
 Db 2315 GTGTACAGGCAATGAGATTAAACAGAAATTTTAAAGGTGTTTCTAGTTCAATGCTTC 2374
 QY 2358 CCGCCGCT-----CCCTGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2399
 Db 2375 CCTTGCCCTGACATTTGTGGCATTTCCACCCCTGCTCTCTCTCTCTCTCTCTCTCTCT 2434
 QY 2400 TGAACACCCAG 2459
 Db 2435 TGAAGACAGCTAGGAG 2493
 QY 2460 AGCGCGGAG 2519
 Db 2494 AGTCCACAGAGCTGAG 2548
 QY 2520 GTCTTGTGCTGTCTCCAG 2579
 Db 2549 ATCTGTGCTGTGCTCTCCAG 2599
 QY 2580 AGAGATAGTCTTCTTCT 2638
 Db 2600 TGAAG-----GTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2654
 QY 2639 CAGTATGCGCTTGTGAG 2690
 Db 2655 TAGTATTGGCTTGTATGAG 2706
 RESULT 15
 AP002776 160903 bp DNA linear HTG 17-JAN-2001
 LOCUS Homo sapiens chromosome 11 clone RP11-126P21 map 11q, WORKING DRAFT
 DEFINITION
 ACCESSION AP002776
 VERSION AP002776.2 GI:12246853
 KEYWORDS HTG: HTGS, PHASE1: HTGS, DRAFT.
 SOURCE Homo sapiens DNA, clone: RP11-126P21.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 160903)
 Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,
 Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
 Homo sapiens 160,903 genomic DNA of 11q
 Published Only in DataBase (2000)
 2 (bases 1 to 160903)
 Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,
 Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
 Direct Submission
 Submitted (10-JUL-2000) Masahisa Hattori, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
 1-7-22 Tsurumi-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 (E-mail: hattori@gs.c.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/

COMMENT

Tel:81-45-503-9111, Fax:81-45-503-9170)
On Jan 16, 2001 this sequence version replaced gi:9188614.
----- Genome Center
Center: RIKEN Genomic Sciences Center(GSC)
Center code: RIKEN
Web site: <http://hgp.gsc.riken.go.jp/>
Contact: hattori@gsc.riken.go.jp
----- Project Information
Center project name: Humpert11
Center clone name: Rp11-126p21
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 14950 bases at least Q40
Consensus quality: 15393 bases at least Q30
Consensus quality: 156144 bases at least Q20
Insert size: 157403; sum-of-contigs
Quality coverage: 8.42x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 36 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs 'N', but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
1 12626 contig of 12626 bp in length
12727 24539 contig of 11813 bp in length
24640 33605 contig of 8966 bp in length
33706 45529 contig of 11824 bp in length
45630 51226 contig of 5597 bp in length
51227 51326 contig of 5597 bp in length
51327 58567 contig of 7241 bp in length
58568 58667 contig of 6821 bp in length
58668 65488 contig of 6727 bp in length
65489 72315 contig of 5882 bp in length
72316 72415 contig of 5716 bp in length
72416 78297 contig of 5882 bp in length
78298 78397 contig of 5882 bp in length
78398 84113 contig of 5716 bp in length
84114 84213 contig of 5716 bp in length
84214 90536 contig of 6323 bp in length
90537 90636 contig of 100 bp
90637 95051 contig of 4415 bp in length
95052 95151 contig of 100 bp
95152 99283 contig of 4132 bp in length
99284 99383 contig of 100 bp
99384 103658 contig of 4275 bp in length
103659 103758 contig of 100 bp
103759 107823 contig of 4065 bp in length
107824 107923 contig of 100 bp
107924 113135 contig of 5312 bp in length
113136 113235 contig of 100 bp
113236 117634 contig of 4399 bp in length
117635 117734 contig of 100 bp
117735 121241 contig of 3507 bp in length
121242 121341 contig of 100 bp
121342 124894 contig of 3553 bp in length
124895 124994 contig of 100 bp
124995 128257 contig of 3263 bp in length
128258 128357 contig of 100 bp
128358 131408 contig of 3051 bp in length
131409 131508 contig of 100 bp
131509 134212 contig of 2704 bp in length
134213 134312 contig of 100 bp
134313 136588 contig of 2276 bp in length
136589 136688 contig of 100 bp
136689 139497 contig of 2809 bp in length
139498 139597 contig of 100 bp
139598 142074 contig of 2477 bp in length
142075 142174 contig of 100 bp
142175 143858 contig of 1684 bp in length
143859 143958 contig of 100 bp
143959 146242 contig of 2284 bp in length
146243 146342 contig of 100 bp
146343 148167 contig of 1825 bp in length
148168 148267 contig of 100 bp
148268 150246 contig of 1979 bp in length
150247 150346 contig of 100 bp
150347 151876 contig of 1530 bp in length
151877 151976 contig of 100 bp
151977 154322 contig of 2246 bp in length
154323 154422 contig of 100 bp
154423 155669 contig of 1247 bp in length
155670 155769 contig of 100 bp
155770 156932 contig of 1163 bp in length
156933 157032 contig of 100 bp
157033 158659 contig of 1627 bp in length
158660 158759 contig of 100 bp
158760 159622 contig of 863 bp in length
159623 159722 contig of 100 bp
159723 160903 contig of 1181 bp in length.

NOTE: This is a 'working draft' sequence. It currently consists of 36 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
1 12627 contig of 12626 bp in length
12627 12726 gap of 100 bp

FEATURES

source

* 12727 24539: contig of 11813 bp in length
* 24540 24639: gap of 100 bp
* 24640 33605: contig of 8966 bp in length
* 33606 33705: gap of 100 bp
* 33706 45529: contig of 11824 bp in length
* 45530 45629: gap of 100 bp
* 45630 51226: contig of 5597 bp in length
* 51227 51326: gap of 100 bp
* 51327 58567: contig of 7241 bp in length
* 58568 58667: gap of 100 bp
* 58668 65488: contig of 6821 bp in length
* 65489 72315: gap of 100 bp
* 72316 72415: gap of 100 bp
* 72416 78297: contig of 5882 bp in length
* 78298 78397: gap of 100 bp
* 78398 84113: contig of 5716 bp in length
* 84114 84213: gap of 100 bp
* 84214 90536: contig of 6323 bp in length
* 90537 90636: gap of 100 bp
* 90637 95051: contig of 4415 bp in length
* 95052 95151: gap of 100 bp
* 95152 99283: contig of 4132 bp in length
* 99284 99383: gap of 100 bp
* 99384 103658: contig of 4275 bp in length
* 103659 103758: gap of 100 bp
* 103759 107823: contig of 4065 bp in length
* 107824 107923: gap of 100 bp
* 107924 113135: contig of 5312 bp in length
* 113136 113235: gap of 100 bp
* 113236 117634: contig of 4399 bp in length
* 117635 117734: gap of 100 bp
* 117735 121241: contig of 3507 bp in length
* 121242 121341: gap of 100 bp
* 121342 124894: contig of 3553 bp in length
* 124895 124994: gap of 100 bp
* 124995 128257: contig of 3263 bp in length
* 128258 128357: gap of 100 bp
* 128358 131408: contig of 3051 bp in length
* 131409 131508: gap of 100 bp
* 131509 134212: contig of 2704 bp in length
* 134213 134312: gap of 100 bp
* 134313 136588: contig of 2276 bp in length
* 136589 136688: gap of 100 bp
* 136689 139497: contig of 2809 bp in length
* 139498 139597: gap of 100 bp
* 139598 142074: contig of 2477 bp in length
* 142075 142174: gap of 100 bp
* 142175 143858: contig of 1684 bp in length
* 143859 143958: gap of 100 bp
* 143959 146242: contig of 2284 bp in length
* 146243 146342: gap of 100 bp
* 146343 148167: contig of 1825 bp in length
* 148168 148267: gap of 100 bp
* 148268 150246: contig of 1979 bp in length
* 150247 150346: gap of 100 bp
* 150347 151876: contig of 1530 bp in length
* 151877 151976: gap of 100 bp
* 151977 154322: contig of 2246 bp in length
* 154323 154422: gap of 100 bp
* 154423 155669: contig of 1247 bp in length
* 155670 155769: gap of 100 bp
* 155770 156932: contig of 1163 bp in length
* 156933 157032: gap of 100 bp
* 157033 158659: contig of 1627 bp in length
* 158660 158759: gap of 100 bp
* 158760 159622: contig of 863 bp in length
* 159623 159722: gap of 100 bp
* 159723 160903: contig of 1181 bp in length.
1. 160903 location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"

Db	241	CATGGAGGGGACATGATGATGACAGACAGGCCACTTCTGAGCCCAACAGAGAGAGGCCCGG	300
Qy	301	AGTGAAGAGAGAGCTCCACGGGGACCAGACAGATTCTGGGGCAAGATGCCAGATGCCAG	360
Db	301	AGTGAAGAGAGAGCTCCACGGGGACCAGACAGACTTCGGGGCAAGGATGCCAGATGCCAG	360
Qy	361	AAGCAGGAGAGAGAGAGGACAGCTGCACTTCATGATGATACAGCTGCTAGAGGCGCAGAGAT	420
Db	361	AAGCAGGAGAGAGAGAGGACAGCACTGCACTTCATGATGATACAGCTGCTAGAGGCGCAGAGAT	420
Qy	421	GACATCCGCGCTGGCAGGCCAGCTGAGAGGACCCCGGGCTCCCGGGCTCGCTACCTGCTG	480
Db	421	GACATCCGCGCTGGCAGGCCAGCTGAGAGGACCCCGGGCTCCCGGGCTCGCTACCTGCTG	480
Qy	481	GTAGATTCTACACGAGAAAGAGAGAGGTCTGAGCCAGATGAGACGGTCTCTGGGCGTG	540
Db	481	GTAGATTCTACACGAGAAAGAGAGAGGTCTGAGCCAGATGAGACGGTCTCTGGGCGTG	540
Qy	541	GATTTTCCTGACAGCAGCTCCCCCAGCTGACACCTTGCGGCTGCTTGCCCTCTGGAGT	600
Db	541	GATTTTCCTGACAGCAGCTCCCCCAGCTGACACCTTGCGGCTGCTTGCCCTCTGGAGT	600
Qy	601	GACACCAGAGTACTACTAGATGATGAGAGCGGGGCTTCAGGCTGACGCTCTGGTGGGCAAGC	660
Db	601	GACACCAGAGTACTACTAGATGATGAGAGCGGGGCTTCAGGCTGACGCTCTGGTGGGCAAGC	660
Qy	661	CGGATCTTCAGGCCATCTCATCCAGACCATGTGGGCCACACTCCAGTATTGCACAA	720
Db	661	CGGATCTTCAGGCCATCTCATCCAGACCATGTGGGCCACACTCCAGTATTGCACAA	720
Qy	721	GCATGTGAGGCAAGCTTATAGGCAAGCGGGCTTTTAAACGGGGTGGCAGTGGCCCTCACTGGGGC	780
Db	721	GCATGTGAGGCAAGCTTATAGGCAAGCGGGCTTTTAAACGGGGTGGCAGTGGCCCTCACTGGGGC	780
Qy	781	AGCCACTACAGAGAGAGACTGGAACCTCCGAACAGAGCTGCTCAATGATGAGAGGCTATG	840
Db	781	AGCCACTACAGAGAGAGACTGGAACCTCCGAACAGAGCTGCTCAATGATGAGAGGCTATG	840
Qy	841	GGCGACCTGGAGTCTGTGGGGCTCCAGACGCCCGAGCCTTGGCGGGTCTCTGAAACAGGAG	900
Db	841	GGCGACCTGGAGTCTGTGGGGCTCCAGAGGCCCGAGCCTTGGCGGGTCTCTGAAACAGGAG	900
Qy	901	CAGATGGAGCAGGCCGATCCGTGCTGAGCTGTGGAAGTGTGGATGTCACTGACCTGGAG	960
Db	901	CAGATGGAGCAGGCCGATCCGTGCTGAGCTGTGGAAGTGTGGATGTCACTGACCTGGAG	960
Qy	961	AGTGTACTTCCAAAGAGATCCGCCAGGCCCTGAGAGCTGGGCGCTCCGCCCTCAG	1020
Db	961	AGTGTACTTCCAAAGAGATCCGCCAGGCCCTGAGAGCTGGGCGCTCCGCCCTCAG	1020
Qy	1021	CAGTACCCTGACTTCATGCAACCAACAGATGCTGCTGCTGCTGAGACAGCGGAGCAAGCC	1080
Db	1021	CAGTACCCTGACTTCATGCAACCAACAGATGCTGCTGCTGCTGAGACAGCGGAGCAAGCC	1080
Qy	1081	TCCCGCATCTTCCCGCACCTCTACCTTGGGCTCAGAGTGAAGCAGCAAACTGGAGAG	1140
Db	1081	TCCCGCATCTTCCCGCACCTCTACCTTGGGCTCAGAGTGAAGCAGCAAACTGGAGAG	1140
Qy	1141	CTGCAAGAGAAACAGGGTCAACCCACATCTTGAACATGGGCCGGGAGATGTCAACTCTTAC	1200
Db	1141	CTGCAAGAGAAACAGGGTCAACCCACATCTTGAACATGGGCCGGGAGATGTCAACTCTTAC	1200
Qy	1201	CTGAGAGGCTTCACTACCAACAATGTGCGGCTCTGGGATGAGAGTGGGCCAGCTGCTG	1260
Db	1201	CTGAGAGGCTTCACTACCAACAATGTGCGGCTCTGGGATGAGAGTGGGCCAGCTGCTG	1260
Qy	1261	CCGCACTGGAAGGAGAGCAGCCGCTTCAATTGAGGCTCAAGAGACAGGGCACCCACAGCTG	1320
Db	1261	CCGCACTGGAAGGAGAGCAGCCGCTTCAATTGAGGCTCAAGAGACAGGGCACCCACAGCTG	1320
Qy	1321	CTGGTCCACTGCAAGATGGGCGTCAAGCCGCTCAGCGGCCACAGTGTGGGCTATGCGATG	1380
Db	1321	CTGGTCCACTGCAAGATGGGCGTCAAGCCGCTCAGCGGCCACAGTGTGGGCTATGCGATG	1380
Qy	1321	CTGGTCCACTGCAAGATGGGCGTCAAGCCGCTCAGCGGCCACAGTGTGGGCTATGCGATG	1380
Db	1321	CTGGTCCACTGCAAGATGGGCGTCAAGCCGCTCAGCGGCCACAGTGTGGGCTATGCGATG	1380

Qy	1381	AAGCAGTACGAATGACAGCCTTGAGAGCAGCCCTGGGCCACATGCAAGAGAGCTCCGGCCATC	1440
Db	1381	AAGCAGTACGAATGACAGCCTTGAGAGCAGCCCTGGGCCACATGCAAGAGAGCTCCGGCCATC	1440
Qy	1441	GGCCGGCCCCAACCCCTGGCTCCCTGGCCAGCTGACGATCTACAGGGGATCTTACGGCC	1500
Db	1441	GGCCGGCCCCAACCCCTGGCTCCCTGGCCAGCTGACGATCTACAGGGGATCTTACGGCC	1500
Qy	1501	AGAACCTGAGGGTGTGGGGAGAGAGGTTGTAGCATGTGAAGAGAGCCAGGAGCCCC	1560
Db	1501	AGAACCTGAGGGTGTGGGGAGAGAGGTTGTAGCATGTGAAGAGAGCCAGGAGCCCC	1560
Qy	1561	GAAAGAGAGGCTTGGGGCCAGCGGCCAGTAAATCAACCTCCGAGGGGTCATGAGGTCATCAG	1620
Db	1561	GAAAGAGAGGCTTGGGGCCAGCGGCCAGTAAATCAACCTCCGAGGGGTCATGAGGTCATCAG	1620
Qy	1621	TCCTTGGAGGCCCTCTTGGAGCTGGAGAGCACTCAGAGACCACTGACATGCCAGAGCT	1680
Db	1621	TCCTTGGAGGCCCTCTTGGAGCTGGAGAGCACTCAGAGAGCACTGACATGCCAGAGAGT	1680
Qy	1681	CTTCTCTTCCCAAGATCTTTCACATGAAGAGCTGTGAGAGCCCTTCCACAGCTTGCAG	1740
Db	1681	CTTCTCTTCCCAAGATCTTTCACATGAAGAGCTGTGAGAGCCCTTCCACAGCTTGCAG	1740
Qy	1741	GACCAAGGAGAGGCAAGCAGTGGTGAACAGGGGGCCTCAGCCTGGCCCTGAAGTCCCGCACTC	1800
Db	1741	GACCAAGGAGAGGCAAGCAGTGGTGAACAGGGGGCCTCAGCCTGGCCCTGAAGTCCCGCACTC	1800
Qy	1801	AGTGTATTACCTCTCAGGGGCAATGCGGTGTGGCCAAACGGAGACCCAGGCGCTTCCAGAGACA	1860
Db	1801	AGTGTATTACCTCTCAGGGGCAATGCGGTGTGGCCAAACGGAGACCCAGGCGCTTCCAGAGACA	1860
Qy	1861	GGAGCAGGGGCAAGGGGCAAGGGGCAAGGCAAGGCAATTTCTCTTACAGCCAGATTCCG	1920
Db	1861	GGAGCAGGGGCAAGGGGCAAGGGGCAAGGCAAGGCAATTTCTCTTACAGCCAGATTCCG	1920
Qy	1921	GAAAGTGTGAGAGCAGGGCCAGCGTGCATGACAGTGAAGAGAGGGCCAGAGCCTTGAGCCTT	1980
Db	1921	GAAAGTGTGAGAGCAGGGCCAGCGTGCATGACAGTGAAGAGAGGGCCAGAGCCTTGAGCCTT	1980
Qy	1981	CACACATGCCCCAAGCTTCCCTTCAACCTGAAGAGAGATCCACACATCTCTTGGAGAAACACCC	2040
Db	1981	CACACATGCCCCAAGCTTCCCTTCAACCTGAAGAGAGATCCACACATCTCTTGGAGAAACACCC	2040
Qy	2041	TCAAGTCTGTTGGCGCACATTCCTCTCAGTCCGCGCCCATACCCGTCATACAGCCTC	2100
Db	2041	TCAAGTCTGTTGGCGCACATTCCTCTCAGTCCGCGCCCATACCCGTCATACAGCCTC	2100
Qy	2101	ACCTCCCAACCCCTGTCACTAAGGCTCTCACCCTGTCTACTATCAAGCTTCACCTCC	2160
Db	2101	ACCTCCCAACCCCTGTCACTAAGGCTCTCACCCTGTCTACTATCAAGCTTCACCTCC	2160
Qy	2161	TACAGCCTTAAAGTCCCAAGGCGCCATGTCTGCTCTGTCCAAAGGGCTCAAGATTTCTAACTGG	2220
Db	2161	TACAGCCTTAAAGTCCCAAGGCGCCATGTCTGCTCTGTCCAAAGGGCTCAAGATTTCTAACTGG	2220
Qy	2221	GATGTGTAGAGGAGCTGAAGTACTTTGGGGGCAACAGACACCTAGTTTATTCTCAA	2280
Db	2221	GATGTGTAGAGGAGCTGAAGTACTTTGGGGGCAACAGACACCTAGTTTATTCTCAA	2280
Qy	2281	CTCTAGCCCTGACACTCACTTGTGGACAGGAATGAAAACAGAGCTTCCCGTGCAGAAAAG	2340
Db	2281	CTCTAGCCCTGACACTCACTTGTGGACAGGAATGAAAACAGAGCTTCCCGTGCAGAAAAG	2340
Qy	2341	GGTACAGCCTCCCAACCCGCGCCCTCTCCGTGACCTCTCTCTCCAGTTTACTTCTCT	2400
Db	2341	GGTACAGCCTCCCAACCCGCGCCCTCTCCGTGACCTCTCTCTCTCCAGTTTACTTCTCT	2400
Qy	2401	GGAACAGGCCAGGCGCAGGCAACAGTGGCCCCCAAAGGCAAGGAGAGATCTCAAGGCCCCA	2460
Db	2401	GGAACAGGCCAGGCGCAGGCAACAGTGGCCCCCAAAGGCAAGGAGAGATCTCAAGGCCCCA	2460

Oy 2461 GCCGGGAGAGCTGGAGGGCTGGAGATCGTTCCATCCATCCACCGGCTCAGG 2520
Db 2461 GCCGGGAGAGCTGGAGGGCTGGAGATCGTTCCATCCATCCACCGGCTCAGG 2520
Oy 2521 TCTTGTGCTGTGCTCCCGAGACCTCTGTGACACCCAGCCAGATCAGAGGCCA 2580
Db 2521 TCTTGTGCTGTGCTCCCGAGACCTCTGTGACACCCAGCCAGATCAGAGGCCA 2580
Oy 2581 GAGATAGCTTCTTTTCTTCCCTTCTGCGCTCTGCGTGTAGTGTTCATAGCCTTACA 2640
Db 2581 GAGATAGCTTCTTTTCTTCCCTTCTGCGCTCTGCGTGTAGTGTTCATAGCCTTACA 2640
Oy 2641 GATCTGCGCTTGTACTGAGAAATTAACACATTTTCATTAATAAAAAAAAAAAAAA 2700
Db 2641 GATCTGCGCTTGTACTGAGAAATTAACACATTTTCATTAATAAAAAAAAAAAAAA 2700
Oy 2701 AAAA 2704
Db 2701 AAAA 2704

RESULT 2
US-09-761-640-7
: Sequence 7, Application US/09761640
: Patent No. US20020137042A1
: GENERAL INFORMATION:
: APPLICANT: WEI, Ming-Hui et al
: TITLE OF INVENTION: ISOLATED HUMAN PHOSPHATASE PROTEINS,
: TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHATASE PROTEINS,
: TITLE OF INVENTION: AND USES THEREOF
: FILE REFERENCE: CL000964-CIP
: CURRENT APPLICATION NUMBER: US/09/761, 640
: CURRENT FILING DATE: 2001-01-18
: NUMBER OF SEQ ID NOS: 10
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 7
: LENGTH: 2704
: TYPE: DNA
: ORGANISM: Human
US-09-761-640-7

Query Match 100.0%; Score 2704; DB 10; Length 2704;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2704; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GCTCTCTCTGCTCTGCGGCTCCAGACTGTCCGCGGGGTTGAGGAAAGGGCCGTGCC 60
Db 1 GCTCTCTCTGCTCTGCGGCTCCAGACTGTCCGCGGGGTTGAGGAAAGGGCCGTGCC 60
Oy 61 CGGTGCCAGCCAGAGTGTCTCGCGGCTCGCTCATGCGCTGTGACAGTGAAGCGTTTCG 120
Db 61 CGGTGCCAGCCAGAGTGTCTCGCGGCTCGCTCATGCGCTGTGACAGTGAAGCGTTTCG 120
Oy 121 CCCCCGAGAGCGGCGCTCCACGCGCTGCGGCGCTTGAGGACGAGCGGTCCAGCGAAAG 180
Db 121 CCCCCGAGAGCGGCGCTCCACGCGCTGCGGCGCTTGAGGACGAGCGGTCCAGCGAAAG 180
Oy 181 AGTGACATCCAGCGAAGGACAGAGCTTTCGCGTCCGTGGGGCTGTCTGGAGTGTGAG 240
Db 181 AGTGACATCCAGCGAAGGACAGAGCTTTCGCGTCCGTGGGGCTGTCTGGAGTGTGAG 240
Oy 241 GATGAGAGGAGCAATGATGATGACAGAGGCGCACTTGTGAGCCAAAGAGAGGCCCG 300
Db 241 GATGAGAGGAGCAATGATGATGACAGAGGCGCACTTGTGAGCCAAAGAGAGGCCCG 300
Oy 301 AGTGAGAGAGAGCTCCACGCGGAGCAGACAGACTTTCGCGCAAGATGCCAGATCCCGAG 360
Db 301 AGTGAGAGAGAGCTCCACGCGGAGCAGACAGACTTTCGCGCAAGATGCCAGATCCCGAG 360
Oy 361 AAGCAGAGAGAGCAGAGAGCAGACCTGACATGATGATGACAGAGCGCGGAGAT 420
Db 361 AAGCAGAGAGAGCAGAGAGCAGACCTGACATGATGATGACAGAGCGCGGAGAT 420

Oy 421 GACATCCGCTTGAGAGAGCCAGCTGAGAGCACCAGGCTCTCCCGGCTCGCTACCTGTG 480
Db 421 GACATCCGCTTGAGAGAGCCAGCTGAGAGCACCAGGCTCTCCCGGCTCGCTACCTGTG 480
Oy 481 GTAGTTTCTACACGAGAAAGAGAGAGCTGTGAGCTCAGAGATGAGACGGTCTCTGGCGCTG 540
Db 481 GTAGTTTCTACACGAGAAAGAGAGAGCTGTGAGCTCAGAGATGAGACGGTCTCTGGCGCTG 540
Oy 541 GATTTCCTGAGAGAGAGCTCCCCAGCTGACCCAGCTGGGCGCTGTGGCTCTGGAGAT 600
Db 541 GATTTCCTGAGAGAGAGCTCCCCAGCTGACCCAGCTGGGCGCTGTGGCTCTGGAGAT 600
Oy 601 GACACCCAGCTGTACTTGTAGATGAGAGCGGGGCTTCAGCGTGTGCTGTGGGCAAAAGC 660
Db 601 GACACCCAGCTGTACTTGTAGATGAGAGCGGGGCTTCAGCGTGTGCTGTGGGCAAAAGC 660
Oy 661 CGGATCTTCAAGCCCATCTCCATCCAGACCATGTGGGCCACACTCCAGGTATTGCACCAA 720
Db 661 CGGATCTTCAAGCCCATCTCCATCCAGACCATGTGGGCCACACTCCAGGTATTGCACCAA 720
Oy 721 GCATGTGAGGCGAGCTCTGAGGCGAGGCGCTTGTACCGGGGTGAGTGGCCCTACCTGGGGCC 780
Db 721 GCATGTGAGGCGAGCTCTGAGGCGAGGCGCTTGTACCGGGGTGAGTGGCCCTACCTGGGGCC 780
Oy 781 AGCCACTACAGAGAGAGACTCCGAACAGAGCTGCTCAATAGATGAGTGAAGGCTATG 840
Db 781 AGCCACTACAGAGAGAGACTCCGAACAGAGCTGCTCAATAGATGAGTGAAGGCTATG 840
Oy 841 GCCGACCTGAGAGTCTCTGCGGCGCTCCAGCGCGAGCGCTGGCGGGTCTCCAGAAACAGAG 900
Db 841 GCCGACCTGAGAGTCTCTGCGGCGCTCCAGCGCGAGCGCTGGCGGGTCTCCAGAAACAGAG 900
Oy 901 CAGATGTGAGGCGAGCTCCGCTGAGCGTGAAGGTTGAGTGAAGTGAAGTGAAGTGAAG 960
Db 901 CAGATGTGAGGCGAGCTCCGCTGAGCGTGAAGGTTGAGTGAAGTGAAGTGAAGTGAAG 960
Oy 961 AGTGTCACTTCCAAAGAGATCCGCGAGGCTGTGAGCTGCGGCTCGGCTCGGCTCGGCTCGG 1020
Db 961 AGTGTCACTTCCAAAGAGATCCGCGAGGCTGTGAGCTGCGGCTCGGCTCGGCTCGGCTCGG 1020
Oy 1021 CAGTACCGTGTGACTTATGAGACACAGATGCTGCTGTGTGTGTGTGTGTGTGTGTGTGT 1080
Db 1021 CAGTACCGTGTGACTTATGAGACACAGATGCTGCTGTGTGTGTGTGTGTGTGTGTGTGT 1080
Oy 1081 TCCCGCATCTTCCCGACCTTACCTGAGGCTGAGAGTGAAGGCAAGCAGCAACCTGAGAGG 1140
Db 1081 TCCCGCATCTTCCCGACCTTACCTGAGGCTGAGAGTGAAGGCAAGCAGCAACCTGAGAGG 1140
Oy 1141 CTGCAGAGGAGACAGGGTCAACCATCTTGAACATGAGGCGCGGAGATTGACAACTTCTTAC 1200
Db 1141 CTGCAGAGGAGACAGGGTCAACCATCTTGAACATGAGGCGCGGAGATTGACAACTTCTTAC 1200
Oy 1201 CTTGAGCGCTTACCTTACCAATGTGCGCTCTGTGGATGAGAGATCGGCCAGCTGCTG 1260
Db 1201 CTTGAGCGCTTACCTTACCAATGTGCGCTCTGTGGATGAGAGATCGGCCAGCTGCTG 1260
Oy 1261 CCGCACTGGAAGGAGAGCAGCAGCTTCAFTGAGGCTGGAAGGACACAGGACCCACAGCG 1320
Db 1261 CCGCACTGGAAGGAGAGCAGCAGCTTCAFTGAGGCTGGAAGGACACAGGACCCACAGCG 1320
Oy 1321 CTGTGTCACTGTCAAGATGAGGCTGAGCGCTGAGCGGCTGAGCGGCTGAGCTGTATTCATG 1380
Db 1321 CTGTGTCACTGTCAAGATGAGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCTGTATTCATG 1380
Oy 1381 AAGCAGTACGAATGCAAGCTGTGAGCAGGCTGTGCGCCACAGTGTGAGAGAGTCCGCGCATC 1440
Db 1381 AAGCAGTACGAATGCAAGCTGTGAGCAGGCTGTGCGCCACAGTGTGAGAGAGTCCGCGCATC 1440
Oy 1441 GCGCGCCCAACCCGCGCTTCTGCGCAGCTGCAAGATCTACAGGGATCTGTGAGGCGC 1500
Db 1441 GCGCGCCCAACCCGCGCTTCTGCGCAGCTGCAAGATCTACAGGGATCTGTGAGGCGC 1500
Oy 1501 AGAACCTGAGGGTGTGTGGGAGAGAGGTTGTAGGATGGAAGAGAGCAGGAGAGCC 1560

|||||
 Db 1501 AGAAGCTAGGGTGGTGGGAGAGAGAGTTGTAGCATGGAAGAGAGGACAGGACCC 1560
 QY 1561 GAAAGAGGCGTGGGCGACGCGCATTAACCTCCGAGGGGATCATAGGTTCCATVAG 1620
 Db 1561 GAAAGAGGCGTGGGCGACGCGCATTAACCTCCGAGGGGATCATAGGTTCCATVAG 1620
 QY 1621 TCTTCTGAGAGGCGTGGGCGACGCGCATTAACCTCCGAGGGGATCATAGGTTCCATVAG 1680
 Db 1621 TCTTCTGAGAGGCGTGGGCGACGCGCATTAACCTCCGAGGGGATCATAGGTTCCATVAG 1680
 QY 1681 CTTTCTTCCACGAGTCTTCAATGAGAGGCTTGCAGGCTTCCACAGCTTGGCAAG 1740
 Db 1681 CTTTCTTCCACGAGTCTTCAATGAGAGGCTTGCAGGCTTCCACAGCTTGGCAAG 1740
 QY 1741 GACCAAGGAGGCGACGAGTGGACAGAGGGGCGCTCAGCTGCGCTGAAGTCCGCGAGTC 1800
 Db 1741 GACCAAGGAGGCGACGAGTGGACAGAGGGGCGCTCAGCTGCGCTGAAGTCCGCGAGTC 1800
 QY 1801 AGTGTACCTTCAGAGGAGTGGCGTGGCCAAACCGGAGCCCTTCCAGAGAGA 1860
 Db 1801 AGTGTACCTTCAGAGGAGTGGCGTGGCCAAACCGGAGCCCTTCCAGAGAGA 1860
 QY 1861 GGAGCAGGGGCGAGGGGCGAGGGGAGAGAGGCGCTGCAATTTCTTACGCGGATTCG 1920
 Db 1861 GGAGCAGGGGCGAGGGGCGAGGGGAGAGAGGCGCTGCAATTTCTTACGCGGATTCG 1920
 QY 1921 GAAGGTGTGAGACAGAGGCGAGGCTGATACAGTGAAGAGAGGCGAGGCGTGAAGCCT 1980
 Db 1921 GAAGGTGTGAGACAGAGGCGAGGCTGATACAGTGAAGAGAGGCGAGGCGTGAAGCCT 1980
 QY 1981 CACACATGCCACAGCTCCCTGACACTGAAGAGATCCACAACTCTTGGAGAACACC 2040
 Db 1981 CACACATGCCACAGCTCCCTGACACTGAAGAGATCCACAACTCTTGGAGAACACC 2040
 QY 2041 TCACGCTTGTGGCGACATTTCTCAGGCTCCGCGCCATACCGCTGACTACAGCCTC 2100
 Db 2041 TCACGCTTGTGGCGACATTTCTCAGGCTCCGCGCCATACCGCTGACTACAGCCTC 2100
 QY 2101 ACCTCCACCCCTGTGACTACAGGCGCTCACCTCCACCCCTGTGACTACAGCCTCACCTCC 2160
 Db 2101 ACCTCCACCCCTGTGACTACAGGCGCTCACCTCCACCCCTGTGACTACAGCCTCACCTCC 2160
 QY 2161 TACAGCTTAAGTCCAGGCGCATGTGCTGCTGCTCCAGAGGCTCAAGACTTTTAACTGG 2220
 Db 2161 TACAGCTTAAGTCCAGGCGCATGTGCTGCTGCTCCAGAGGCTCAAGACTTTTAACTGG 2220
 QY 2221 GATGTGTGAGAGGAGTGAAGTACCTTTGGGGGCAACAGCACCTAGTTTCTCA 2280
 Db 2221 GATGTGTGAGAGGAGTGAAGTACCTTTGGGGGCAACAGCACCTAGTTTCTCA 2280
 QY 2281 CTTTACCCCTGCACACTACCTGTGGCAGGAAATGAAGAGAGCTTCCGTCAAAAAG 2340
 Db 2281 CTTTACCCCTGCACACTACCTGTGGCAGGAAATGAAGAGAGCTTCCGTCAAAAAG 2340
 QY 2341 GGTACAGGCTCCAGCCCGCCCTCCTGCACTGCTCTCTCTCCAGTTTCAATTCCT 2400
 Db 2341 GGTACAGGCTCCAGCCCGCCCTCCTGCACTGCTCTCTCTCCAGTTTCAATTCCT 2400
 QY 2401 GGAACACGAGGCGAGGCAACAGAGTGGGCGCCCAAGAGGCGAGAGTCTCGAGGCCCA 2460
 Db 2401 GGAACACGAGGCGAGGCAACAGAGTGGGCGCCCAAGAGGCGAGAGTCTCGAGGCCCA 2460
 QY 2461 GCGGCGGAGGCTGGAAGGCTGGAGATGCTTCCCTCATTCACCTCCACGCGTCCAGG 2520
 Db 2461 GCGGCGGAGGCTGGAAGGCTGGAGATGCTTCCCTCATTCACCTCCACGCGTCCAGG 2520
 QY 2521 TCTTGTGCTGTGCCAGACTCTGTGTGACACACGCGAGATCAGAGGCGACAGGCCA 2580
 Db 2521 TCTTGTGCTGTGCCAGACTCTGTGTGACACACGCGAGATCAGAGGCGACAGGCCA 2580
 QY 2581 GAATAGTCTTCTTTTGTCTTGTGGCTCTGGCTAGTCACTTTTCTATACCTTACA 2640
 Db 2581 GAATAGTCTTCTTTTGTCTTGTGGCTCTGGCTAGTCACTTTTCTATACCTTACA 2640

Db 2581 GAGATAGTCTTCTTTTGTCTTGTGGCTCTGGCTAGTCACTTTTCTATACCTTACA 2640
 QY 2641 GATATGCGCTTTGTTACTAGAAATATAACATTTTCAATTAATAAAAAAAAAAAAAA 2700
 Db 2641 GATATGCGCTTTGTTACTAGAAATATAACATTTTCAATTAATAAAAAAAAAAAAAA 2700
 QY 2701 AAAA 2704
 Db 2701 AAAA 2704
 RESULT 3
 US-09-761-640-2
 ; Sequence 2, Application US/09761640
 ; Patent No. US20020137042A1
 ; GENERAL INFORMATION:
 ; APPLICANT: WEI, Ming-Hui et al
 ; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHATASE PROTEINS,
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHATASE PROTEINS,
 ; TITLE OF INVENTION: AND USES THEREOF
 ; FILE REFERENCE: C1000964-CIP
 ; CURRENT APPLICATION NUMBER: US/09/761,640
 ; CURRENT FILING DATE: 2001-01-18
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 2852
 ; TYPE: DNA
 ; ORGANISM: Human
 US-09-761-640-2
 Query Match 94.6% Score 2558; DB 10; Length 2852;
 Best Local Similarity 96.5% Pred. No. 0;
 Matches 2666; Conservative 0; Mismatches 0; Indels 98; Gaps 1;
 QY 39 GGTGAGGAGAGGCGCGTGGCGGCGTGCAGCCAGAGTGTCTGCGGCGTGCCTCATGCG 98
 Db 2 GGTGAGGAGAGGCGCGTGGCGGCGTGCAGCCAGAGTGTCTGCGGCGTGCCTCATGCG 61
 QY 99 CCTGTGACAGTGAAGCGTTGCGCCCGCGGCGAGCGGCGCTTCCAGCGCCGTGGGCGCTTG 158
 Db 62 CCGTGTGACAGTGAAGCGTTGCGCCCGCGGCGAGCGGCGCTTCCAGCGCCGTGGGCGCTTG 121
 QY 159 GAGCCAGGGGCGTCCAGGAGAGTGCAGTCCAGGAGAGGAGAGGCTTGGGCGTCCG 218
 Db 122 GAGCCAGGGGCGTCCAGGAGAGTGCAGTCCAGGAGAGGAGAGGCTTGGGCGTCCG 181
 QY 219 TGGGCGTGTCTGGGACTCAGAGATGAGGAGGAGCAATGATGATGAGCAGAGGCCAGTTTC 278
 Db 182 TGGGCGTGTCTGGGACTCAGAGATGAGGAGGAGCAATGATGATGAGCAGAGGCCAGTTTC 241
 QY 279 TGAGCCACAGAGAGGCGCCGAGTGAAGAGAGCTCCAGGGGAGACAGAGACTTGGG 338
 Db 242 TGAGCCACAGAGAGGCGCCGAGTGAAGAGAGCTCCAGGGGAGACAGAGACTTGGG 301
 QY 339 GCAAGATATCCAGAGTCCCAAGAGAGAGAGAGAGAGAGAGTCCAGCTCATGATGAT 398
 Db 302 GCAAGATATCCAGAGTCCCAAGAGAGAGAGAGAGAGAGAGTCCAGCTCATGATGAT 361
 QY 399 ACAGTGTGAGGCGCAGAGATGATCCGCTGAGAGCCAGCTGAGAGGAGCCAGCCGCGC 458
 Db 362 ACAGTGTGAGGCGCAGAGATGATCCGCTGAGAGCCAGCTGAGAGGAGCCAGCCGCGC 421
 QY 459 TCCCGGCGTCCGTTACTCTGTGTAATTTCTACAGAGAGAGAGAGTCTGAGGCCAGGA 518
 Db 422 TCCCGGCGTCCGTTACTCTGTGTAATTTCTACAGAGAGAGAGAGTCTGAGGCCAGGA 481
 QY 519 TGAGAGGTCCTCTGAGGCGTGAATTTCCCTGACAGACACTCCCAAGTGCACCTGGG 578
 Db 482 TGAGAGGTCCTCTGAGGCGTGAATTTCCCTGACAGACACTCCCAAGTGCACCTGGG 541
 QY 579 CCGGTCCTGCGCTGTGAGTACACAGGAGTGTACTTGAATGAGAGGAGGCGCTTACG 638
 Db 579 CCGGTCCTGCGCTGTGAGTACACAGGAGTGTACTTGAATGAGAGGAGGCGCTTACG 638

Dh 542 CTTGCTTTGCCCCCTCTGTGAGTGACACCCAGGTGACTTAGATGAGAGCGGGGCTTCAG 601
Oy 639 GGTGACGCTGTGGTGGGCAAGCCGGATCTTCAAGCCCATCTTCATCCAGACCATGTGGGC 698
Dh 602 GGTGACGCTGTGGTGGGCAAGCCGGATCTTCAAGCCCATCTTCATCCAGACCATGTGGGC 661
Oy 699 CACACTCCAGGTATTTGCACCAAGCATGTGAGGACAGCTTAGGACAGCGGCTTTGACCGG 758
Dh 662 CACACTCCAGGTATTTGCACCAAGCATGTGAGGACAGCTTAGGACAGCGGCTTTGACCGG 721
Oy 759 TGGCAGTCCCTCCACCTTGGGCGCAGCCTACTACAGAGAGACTGAACTCCGAACAGAGCTG 818
Dh 722 TGGCAGTCCCTCCACCTTGGGCGCAGCCTACTACAGAGAGACTGAACTCCGAACAGAGCTG 781
Oy 819 CCTCAATGATGAGGAGGCTATGAGCCGACCTGGAGTCTGTGGGCGCTCCAGCGCCGAGCC 878
Dh 782 CCTCAATGATGAGGAGGCTATGAGCCGACCTGGAGTCTGTGGGCGCTCCAGCGCCGAGCC 841
Oy 879 TGGCGGCTCTCAGAACAGAGAGAGATGAGAGCGAGTCCGTGCTGAGCTGTGAAAGT 938
Dh 842 TGGCGGCTCTCAGAACAGAGAGAGATGAGAGCGAGTCCGTGCTGAGCTGTGAAAGT 901
Oy 939 GTTGGATGCTGAGTGACCTGAGAGAGTGTCACTTCCAAAGAGATCCCGCCAGGCTCTGGAGCT 998
Dh 902 GTTGGATGCTGAGTGACCTGAGAGAGTGTCACTTCCAAAGAGATCCCGCCAGGCTCTGGAGCT 961
Oy 999 GCGCTGTGGGGCTCCCGCTCCAGACAGTACCTGACTTCACTGCACACACAGATGCTGCTGCT 1058
Dh 962 GCGCTGTGGGGCTCCCGCTCCAGACAGTACCTGACTTCACTGCACACACAGATGCTGCTGCT 1021
Oy 1059 GGTGCGACAGCGGGAGCGAGCCTCCGAGTCTTCCCGACCTCTACCTGCGGCTGAGAGTG 1118
Dh 1022 GGTGCGACAGCGGGAGCGAGCCTCCGAGTCTTCCCGACCTCTACCTGCGGCTGAGAGTG 1081
Oy 1119 GAAGCGACAAACCTGGAGAGAGCTGACAGAGAACAGGGGTCAACCCACATCTTGAATATGGC 1178
Dh 1082 GAAGCGACAAACCTGGAGAGAGCTGACAGAGAACAGGGGTCAACCCACATCTTGAATATGGC 1141
Oy 1179 CCGGAGATTGTGACAACTTCTACCTCTGAGCGCTTCACTACCAATGTGCGGCTTGCGGA 1238
Dh 1142 CCGGAGATTGTGACAACTTCTACCTCTGAGCGCTTCACTACCAATGTGCGGCTTGCGGA 1201
Oy 1239 TGAGGAGTCCGCGCAGCTGTGCGCAGCTGGAAGAGAGCGACGCGCTTTCATTTGAGGCTGC 1298
Dh 1202 TGAGGAGTCCGCGCAGCTGTGCGCAGCTGGAAGAGAGCGACGCGCTTTCATTTGAGGCTGC 1261
Oy 1299 AAGAGCAAGGGGACCGCAGCTGTGCTGTCACATGCAAGATGGGCGCTACAGCCCTGACGGGC 1358
Dh 1262 AAGAGCAAGGGGACCGCAGCTGTGCTGTCACATGCAAGATGGGCGCTACAGCCCTGACGGGC 1321
Oy 1359 CACAGTGTGGCTATGTGCATGTAAGCATAGCAATGCAAGCTGAGAGAGCGCCCTGCGCA 1418
Dh 1322 CACAGTGTGGCTATGTGCATGTAAGCATAGCAATGCAAGCTGAGAGAGCGCCCTGCGCA 1381
Oy 1419 CGTGAGAGAGCTCCGGCCCATGTGCCGCCCAACCTTGCTTCTGTGCGCAGCTGACAGAT 1478
Dh 1382 CGTGAGAGAGCTCCGGCCCATGTGCCGCCCAACCTTGCTTCTGTGCGCAGCTGACAGAT 1441
Oy 1479 CTACCAGGGGCACTCTGAGG----- 1497
Dh 1442 CTACCAGGGGCACTCTGAGGCGCCAGCCGCGACAGCCATGTCTGAGGAGAGAAAGTGGGTGG 1501
Oy 1498 -----GCC 1500
Dh 1502 GGTCTCCCGACAGAGGACCAAGCCCTGAAAGTCTCTACACATTTCCACCTTCTTCCGCC 1561
Oy 1501 AGAAGCTTAGGGTGGTGGGAGAGAGAGGTTGTAGGCATGGAAGAGAGCCAGAGCCGCC 1560
Dh 1562 AGAAGCTTAGGGTGGTGGGAGAGAGAGGTTGTAGGCATGGAAGAGAGCCAGAGCCGCC 1621
Oy 1561 GAAAGAAAGAGCTGGGCGACAGCGCCACGTATAAAGCTCCGAGGGGCTCATGAGTCCATCAG 1620
Dh 1622 GAAAGAAAGAGCTGGGCGACAGCGCCACGTATAAAGCTCCGAGGGGCTCATGAGTCCATCAG 1681

Oy 1621 TCTTCTGGAGCCCTCTCTTGAGAGCTGGAGAGACCTCAGAGACCCAGTGAATGACAGAGT 1680
Dh 1682 TCTTCTGGAGCCCTCTCTTGAGAGCTGGAGAGACCTCAGAGACCCAGTGAATGACAGAGT 1741
Oy 1681 CTTCCTTCTCCAGAGTCTTCAATGAAGAGCCTCTGACAGCCCTTCCACAGCTTGGCAAG 1740
Dh 1742 CTTCCTTCTCCAGAGTCTTCAATGAAGAGCCTCTGACAGCCCTTCCACAGCTTGGCAAG 1801
Oy 1741 GAGCAAGGGAGGCCAGAGAGTGGACAGGGGGCTCAAGCTTGCCCTGGAAGTCCCGCAGTGC 1800
Dh 1802 GAGCAAGGGAGGCCAGAGAGTGGACAGGGGGCTCAAGCTTGCCCTGGAAGTCCCGCAGTGC 1861
Oy 1801 AGTGGTTACCTCTCAGAGGAGTGGCGGTGGTGGACCAACCGGAGCCAGGCTTCCAGAGAGA 1860
Dh 1862 AGTGGTTACCTCTCAGAGGAGTGGCGGTGGTGGACCAACCGGAGCCAGGCTTCCAGAGAGA 1921
Oy 1861 GGAGCAGGGGCGAGGGGCGAGGGGCGAGAGAGCCGTGCATTTCTCTACGCCAGGTTCCG 1920
Dh 1922 GGAGCAGGGGCGAGGGGCGAGGGGCGAGAGAGCCGTGCATTTCTCTACGCCAGGTTCCG 1981
Oy 1921 GAGGTTGGTGAACAGGGCCAGCGTGCATGACATGAGAGAGGGGCGAGGCGCTGAGCCCT 1980
Dh 1982 GAGGTTGGTGAACAGGGCCAGCGTGCATGACATGAGAGAGGGGCGAGGCGCTGAGCCCT 2041
Oy 1981 CACAGATGGCCACGGCTCCCTGACACTGAAAGAGATTCACAACTCCTTGGAGAAACACCC 2040
Dh 2042 CACAGATGGCCACGGCTCCCTGACACTGAAAGAGATTCACAACTCCTTGGAGAAACACCC 2101
Oy 2041 TCAGCTCTGTTGCGCACACATCTCTCAGTCCGCGCCCATACCCGTCAACAGCCTC 2100
Dh 2102 TCAGCTCTGTTGCGCACACATCTCTCAGTCCGCGCCCATACCCGTCAACAGCCTC 2161
Oy 2101 ACCTCCACCCCTGTGCACAGGCGCTCACCTCCACCCCTGTACATACAGCCTCACCTCC 2160
Dh 2162 ACCTCCACCCCTGTGCACAGGCGCTCACCTCCACCCCTGTACATACAGCCTCACCTCC 2221
Oy 2161 TACAGCCTTAAAGTCCACAGGCCATGTGCTGCTGCAAGAGGGCTCAAGACTTCTTAACTGG 2220
Dh 2222 TACAGCCTTAAAGTCCACAGGCCATGTGCTGCTGCAAGAGGGCTCAAGACTTCTTAACTGG 2281
Oy 2221 GATGTGTAGAGGAGCTGAAGGTACCTTTGGGGGCAACAGCAGCCTTGAATTTCTCAA 2280
Dh 2282 GATGTGTAGAGGAGCTGAAGGTACCTTTGGGGGCAACAGCAGCCTTGAATTTCTCAA 2341
Oy 2281 CTTGAGCCCTGACACACTACCTGTGGCAGGGAATGAAAAAGAGCTTCCCGTGCAGAAAG 2340
Dh 2342 CTTGAGCCCTGACACACTACCTGTGGCAGGGAATGAAAAAGAGCTTCCCGTGCAGAAAG 2401
Oy 2341 GGTACAGGCTCCACAGCCCGCCGCTCCCTGACACTCGTGTCTCCAGAGTTCAATTCCT 2400
Dh 2402 GGTACAGGCTCCACAGCCCGCCGCTCCCTGACACTCGTGTCTCCAGAGTTCAATTCCT 2461
Oy 2401 GGAACAGGCGAGGCCAGGCAACCACTGAGGCGCCCAAGAGGAGGAGAGATCTCAGGCCCA 2460
Dh 2462 GGAACAGGCGAGGCCAGGCAACCACTGAGGCGCCCAAGAGGAGGAGAGATCTCAGGCCCA 2521
Oy 2461 GCGCGGGAGAGGCTGGAAGGGGTGGAGATGCGTTCCTCATCCAGCTTCCACCGGTGCAGG 2520
Dh 2522 GCGCGGGAGAGGCTGGAAGGGGTGGAGATGCGTTCCTCATCCAGCTTCCACCGGTGCAGG 2581
Oy 2521 TCTTTGCTGCTGTCCCGACAGACTCTGTGTGACACAGCGCAGATCCAGAGGCCAGGCCA 2580
Dh 2582 TCTTTGCTGCTGTCCCGACAGACTCTGTGTGACACAGCGCAGATCCAGAGGCCAGGCCA 2641
Oy 2581 GAGATAGTCTTTTCTTTTGTCTTCTGTGCGCTCTGTGCTAGTCAATTAAGCCTTACA 2640
Dh 2642 GAGATAGTCTTTTCTTTTGTCTTCTGTGCGCTCTGTGCTAGTCAATTAAGCCTTACA 2701
Oy 2641 GTATCTGGCTTTGTACTGAGAAATTAACATTTTTCATTAATAAAAAAAAAAAAAAAAAA 2700
Dh 2702 GTATCTGGCTTTGTACTGAGAAATTAACATTTTTCATTAATAAAAAAAAAAAAAAAAAA 2761

QY 934 AAGTGTGATGTCAGTACCTGAGAGTGTCTTCCAAAGAGATCCGCGAGGCTGTG 993
| | | | |
Db 841 AAGTGTGATGTCAGTACCTGAGAGTGTCTTCCAAAGAGATCCGCGAGGCTGTG 900
| | | | |
QY 994 GAGTGTGCTGAGGCTGCTCCCTCCAGAGTACCTGATTCATGCACCAACAGATGCTG 1053
| | | | |
Db 901 GAGTGTGCTGAGGCTGCTCCCTCCAGAGTACCTGATTCATGCACCAACAGATGCTG 960
| | | | |
QY 1054 CTGCTGTGGCAGGAGGAGCGAGCTCCCGATCTCCCACTCCAGCTTACCTGGGCTCA 1113
| | | | |
Db 961 CTGCTGTGGCAGGAGGAGCGAGCTCCCGATCTCCCACTCCAGCTTACCTGGGCTCA 1020
| | | | |
QY 1114 GAGTGTGAGCAGCAAACTGTGAGAGTGTGAGAGAGAGAGGAGTACCCACATCTTTGAA 1173
| | | | |
Db 1021 GAGTGTGAGCAGCAAACTGTGAGAGTGTGAGAGAGAGAGGAGTACCCACATCTTTGAA 1080
| | | | |
QY 1174 ATGGCCCGGAGATTTACAACTTCTTACCTGAGCGCTTACCTACCAATGTGGCCTTC 1233
| | | | |
Db 1081 ATGGCCCGGAGATTTGACAACTTCTTACCTGAGCGCTTACCTACCAATGTGGCCTTC 1140
| | | | |
QY 1234 TGGGATGAGAGTGTGCGCCAGCTGTGCGCAGCTGAGAGAGAGAGCGCTTCAATGAG 1293
| | | | |
Db 1141 TGGGATGAGAGTGTGCGCCAGCTGTGCGCAGCTGAGAGAGAGAGCGCTTCAATGAG 1200
| | | | |
QY 1294 GCTGCAAGAGCAGAGGAGCAGCCAGCTGTGCTGCTCAGAGATGGGCGCTCAGCGCTCA 1353
| | | | |
Db 1201 GCTGCAAGAGCAGAGGAGCAGCCAGCTGTGCTGCTCAGAGATGGGCGCTCAGCGCTCA 1260
| | | | |
QY 1354 GCGGCGACAGTGTGCGCTATGCTATGAGAGAGTACGATGAGAGAGAGAGCGCTTCG 1413
| | | | |
Db 1261 GCGGCGACAGTGTGCGCTATGCTATGAGAGAGTACGATGAGAGAGAGCGCTTCG 1320
| | | | |
QY 1414 CCGCAGCTGAGAGTGTGCGCGCTATGCGCGCGCGCGCGCGCTTCCCTGCGCGAGCTG 1473
| | | | |
Db 1321 CCGCAGCTGAGAGTGTGCGCGCGCTATGCGCGCGCGCGCGCGCTTCCCTGCGCGAGCTG 1380
| | | | |
QY 1474 CAGATCTACAGGCGATCTCTGACG----- 1497
| | | | |
Db 1381 CAGATCTACAGGCGATCTCTGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1440
| | | | |
QY 1498 ----- 1497
| | | | |
Db 1441 GGTGGGCTCTCCAGAGAGAGCAGCCAGCTTCAACCATTTCCACCTCTT 1500
| | | | |
QY 1498 --GCCAGAACTGAGGCTGTGGGAGAGAAAGTGTGAGCATGAGAGAGAGCGAGCA 1555
| | | | |
Db 1501 CCGCAGAACTGAGGCTGTGGGAGAGAAAGTGTGAGCATGAGAGAGAGCGAGCA 1560
| | | | |
QY 1556 GCCCGAGAAAGAGCTGTGGGAGAGCGCGCAGTATAACTCCGAGGGGTCAATGAGGTCC 1615
| | | | |
Db 1561 GCCCGAGAAAGAGCTGTGGGAGAGCGCGCAGTATAACTCCGAGGGGTCAATGAGGTCC 1620
| | | | |
QY 1616 ATCAGCTTCTGAGAGCGCTTGTGAGCTGTGAGAGAGCAGTCAAGAGCAGTGTGAGCA 1675
| | | | |
Db 1621 ATCAGCTTCTGAGAGCGCTTGTGAGCTGTGAGAGAGCAGTCAAGAGCAGTGTGAGCA 1680
| | | | |
QY 1676 GAGGTCTTCTTCCACGAGTCTTACATGAGAGAGCCTGTGAGCGCTTCCACAGCTT 1735
| | | | |
Db 1681 GAGGTCTTCTTCCACGAGTCTTACATGAGAGAGCCTGTGAGCGCTTCCACAGCTT 1740
| | | | |
QY 1736 GCAAGAGCAGAGGAGGAGCAGAGTGTGAGAGAGGAGCGCTCAGCGCTCCCTGAAAGTCCCG 1795
| | | | |
Db 1741 GCAAGAGCAGAGGAGGAGCAGAGTGTGAGAGAGGAGCGCTCAGCGCTCCCTGAAAGTCCCG 1800
| | | | |
QY 1796 CAGTCAAGTGTACCTTCCAGAGCAGTGTGAGTGTGAGCGAGCGAGCGAGCTTCCAG 1855
| | | | |
Db 1801 CAGTCAAGTGTACCTTCCAGAGCAGTGTGAGTGTGAGCGAGCGAGCGAGCTTCCAG 1860
| | | | |
QY 1856 GAGCAGAGCAGGAGGAGGAGCAGAGAGAGAGCGAGCGAGCGAGCTTCCAGTGTGAGCGAG 1915
| | | | |
Db 1861 GAGCAGAGCAGGAGGAGGAGCAGAGAGAGAGCGAGCGAGCGAGCTTCCAGTGTGAGCGAG 1920
| | | | |

QY 1916 TTCCGGAAGTGTGTGAGAGCAGCGCCAGCTGCATGACAGTGTGAGAGAGAGGCGGAGGCTCGA 1975
| | | | |
Db 1921 TTCCGGAAGTGTGTGAGAGCAGCGCCAGCTGCATGACAGTGTGAGAGAGGCGGAGGCTCGA 1980
| | | | |

RESULT 6
US-09-955-732-20
; Sequence 20, Application US/09955732
; Publication No. US20020182203A1
; GENERAL INFORMATION:
; APPLICANT: Luche, Ralf M.
; TITLE OF INVENTION: DSP-15 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.433
; CURRENT APPLICATION NUMBER: US/09/955,732
; CURRENT FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 1416
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-955-732-20

Query Match 52.48; Score 1416; DB 9; Length 1416;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1416; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 94 ATGGCCCTGTGTACAGTGTGAGCGCTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 153
| | | | |
Db 1 ATGGCCCTGTGTACAGTGTGAGCGCTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
| | | | |
QY 154 CCCTGGGAGCAGGCGGTCCAGCGAAGAGTGTGACTCCAGAGGAGAGGCTTTGCGGTG 213
| | | | |
Db 61 CCCTGGGAGCAGGCGGTCCAGCGAAGAGTGTGACTCCAGAGGAGAGGCTTTGCGGTG 120
| | | | |
QY 214 CTCCGTGGGCGTCTCTGAGAGTGTGAGAGTGTGAGAGGAGCAATGATGTGACAGAGCGC 273
| | | | |
Db 121 CTCCGTGGGCGTCTCTGAGAGTGTGAGAGTGTGAGAGGAGCAATGATGTGACAGAGCGC 180
| | | | |
QY 274 AGTTCGAGCCAGAGAGAGAGCGCGAGTGTGAGAGAGAGTGTGAGAGAGAGAGAGAGAG 333
| | | | |
Db 181 AGTTCGAGCCAGAGAGAGAGCGCGAGTGTGAGAGAGAGTGTGAGAGAGAGAGAGAGAG 240
| | | | |
QY 334 TTCCGGAAGTGTGTGAGAGCAGCGCCAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 393
| | | | |
Db 241 TTCCGGAAGTGTGTGAGAGCAGCGCCAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
| | | | |
QY 394 ATGTACAGCTGTGAGAGCGCGAGATGATCCGCTGTGAGAGAGAGAGAGAGAGAGAGAGAG 453
| | | | |
Db 301 ATGTACAGCTGTGAGAGCGCGAGATGATCCGCTGTGAGAGAGAGAGAGAGAGAGAGAGAG 360
| | | | |
QY 454 CGGCGTCCCGCGCTCCGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 513
| | | | |
Db 361 CGGCGTCCCGCGCTCCGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 420
| | | | |
QY 514 CAGATGAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 573
| | | | |
Db 421 CAGATGAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 480
| | | | |
QY 574 CTGGGCTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 633
| | | | |
Db 481 CTGGGCTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 540
| | | | |
QY 634 TTCCAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 693
| | | | |
Db 541 TTCCAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 600
| | | | |
QY 694 TTCCAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 753
| | | | |
Db 601 TTCCAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 660
| | | | |
QY 754 CCGGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 813
| | | | |

|||||
Db 661 CCGGATGGCAGTGGCCCTCACTGGGCGCAGCCACTACAGAGAGAGCTCACTCCGACAG 720
Qy 814 AGCTGCCATATGATGAGAGCGCTATGCGCCAGCTGGAGTCTCTGGCCCTCCAGGCC 873
Db 721 AGCTGCCATATGATGAGAGCGCTATGCGCCAGCTGGAGTCTCTGGCCCTCCAGGCC 780
Qy 874 GAGCTGGCGGCTCTCTCAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 933
Db 781 GAGCTGGCGGCTCTCTCAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Qy 934 AAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 993
Db 841 AAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
Qy 994 GAGCTGGCGGCTCTCTCAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1053
Db 901 GAGCTGGCGGCTCTCTCAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
Qy 1054 CTGCTGGTGGCAGAGCGGAGCGGAGCGCTCCCGCATCTTCCCGCATCTTCCCGCATCT 1113
Db 961 CTGCTGGTGGCAGAGCGGAGCGGAGCGCTCCCGCATCTTCCCGCATCTTCCCGCATCT 1020
Qy 1114 GAGTGAACGAG 1173
Db 1021 GAGTGAACGAG 1080
Qy 1174 ATGGCCCCGGGAGATGAGACACTTCTACCTGAGGCTTACCTACCAATATGTGGCCTC 1233
Db 1081 ATGGCCCCGGGAGATGAGACACTTCTACCTGAGGCTTACCTACCAATATGTGGCCTC 1140
Qy 1234 TGGATGAG 1293
Db 1141 TGGATGAG 1200
Qy 1294 GCTGCAAG 1353
Db 1201 GCTGCAAG 1260
Qy 1354 GCGGCGAAG 1413
Db 1261 GCGGCGAAG 1320
Qy 1414 CGCCACGTGAG 1473
Db 1321 CGCCACGTGAG 1380
Qy 1474 CAGATCTACAG 1509
Db 1381 CAGATCTACAG 1416

RESULT 7
US-09-775-925-1
: Sequence 1, Application US/09775925
: Patent No. US20010049358A1
: GENERAL INFORMATION:
: APPLICANT: Lucbe, Ralf M.
: APPLICANT: Wei, Bo
: TITLE OF INVENTION: DSP-12 AND DSP-13 DUAL-SPECIFICITY
: FILE REFERENCE: PHOSPHATASES
: CURRENT APPLICATION NUMBER: US/09/775,925
: NUMBER OF SEQ ID NOS: 2001-02-01
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 1
: LENGTH: 1949
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-775-925-1

Query Match 12.5%, Score 336.8; DB 10; Length 1949;

Best Local Similarity 59.5%; Pred. No. 2,3e-81;
Matches 591; Conservative 0; Mismatches 397; Indels 6; Gaps 1;

Qy 517 GATGAGAGAGGCTCTCTGAGGCGTGAGATTCCCTGACAGAGAGCTCCCGAGCTGACCTG 576
Db 32 GAGGAGATATCTCTCTGAGGAGGAGCTTTCACATGAGAGAGAGAGAGAGAGAGAGAGAG 91
Qy 577 GAGGAGATATCTCTCTGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 636
Db 92 GAGGAGATATCTCTCTGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 151
Qy 637 AGCGTGAAGTGTGAGGAG 696
Db 152 AGCGTGAAGTGTGAGGAG 211
Qy 697 GCGACACTTCAGAGATTTCCACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 756
Db 212 TCTGCTGAG 271
Qy 757 GATGAGAGAGGCTCTCTGAGGCGTGAGATTCCCTGACAGAGAGCTCCCGAGCTGACCTG 816
Db 272 GATGAGAGAGGCTCTCTGAGGCGTGAGATTCCCTGACAGAGAGCTCCCGAGCTGACCTG 331
Qy 817 TGCTCAATGATGAG 876
Db 332 TGCTCAATGATGAG 391
Qy 877 C-----CTGGGCGGCTCTCTGAGGCGTGAGATTCCCTGACAGAGAGCTCCCGAGCTG 930
Db 392 CTATTTGAG 451
Qy 931 TGGAGAGAGGATGAG 990
Db 452 GAG 511
Qy 991 CTGAG 1050
Db 512 TTAG 571
Qy 1051 CTGCTGCTGAG 1110
Db 572 CTACTTATCTTGGGAG 631
Qy 1111 TCAG 1170
Db 632 TCTGATGAG 691
Qy 1171 AACATGAGGCGGAGAGATGAG 1230
Db 692 AATGTTACAG 751
Qy 1231 CTCTGGGATGAG 1290
Db 752 GCTTACGATGAG 811
Qy 1291 GAGGCTGCAAG 1350
Db 812 AACAG 871
Qy 1351 TCAGGCGGAG 1410
Db 872 TCGGCTCCACAGATGATGAG 931
Qy 1411 CTGGCGGAG 1470
Db 932 TATACATATGTAAG 991
Qy 1471 CTGAGATCTACAG 1504
Db 992 CTGCTGAGATGAG 1025

Query Match 12.5%, Score 336.8; DB 10; Length 1949;

RESULT 8

US-09-775-925-5
: Sequence 5, Application US/09775925
: Patent No. US20010049358A1
: GENERAL INFORMATION:
: APPLICANT: Lucche, Ralf M.
: APPLICANT: Wei, Bo
: TITLE OF INVENTION: DSP-12 AND DSP-13 DUAL-SPECIFICITY
: TITLE OF INVENTION: PHOSPHATASES
: FILE REFERENCE: 200125.420
: CURRENT APPLICATION NUMBER: US/09/775.925
: CURRENT FILING DATE: 2001-02-01
: NUMBER OF SEQ ID NOS: 33
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 5
: LENGTH: 1711
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-775-925-5

Query Match 11.6%; Score 313.4; DB 10; Length 1711;
Best Local Similarity 55.4%; Pred. No. 4.8e-75;
Matches 655; Conservative 0; Mismatches 516; Indels 12; Gaps 2;

QY 379 CAGCAGCTGCACCTGATGATGAGTGGCGGCGGAGATGACATCGGCTGAGCGC 438
DB 348 CAGCTGTCAGAGCATGTCATTTACTCCGCCAGAGACACATCAGGCTGCTGTA 407
QY 439 CAGCTGAGGAGCAGCCCGGCTCCCGCTCCGCTACCTGCTGATGTTCTACAGAGAA 498
DB 408 AGACTGGAAAGTACTTACAGAAATGGAACAGCCTATATGATGTTGATCACTAATGTT 467
QY 499 GGAGAGGCTGAGGCGGAGAGTGGTCTCTGCGGCGGTGATTTCCCTACAGAGCAG 558
DB 468 AGACAAG-----ACACTGAAAGAAAGCATCGTCTAGGATGATTTCTCTATATGAC 521
QY 559 TCCCCAGCTGCACCTGGGCTGCTGCTGCTGCTGAGAGTGAACCCAGGTGACTTA 618
DB 522 ACTGACACTGTGACATGAGGCTTATGTTGCTCTGCTGAGGAGCAGCCTAATTCATTGG 581
QY 619 GATGAGAGCGGGGCTGACGCTGACGCTGTGTTGGGCAAGCGGATCTTCAAGCCATC 678
DB 582 GATGCTGATGATGGGCTGCTGATGATGAGAGGATTAACAGAGTTACATATTAACCTGTA 641
QY 679 TCCATCCAGACCATGTTGGGCCACTCCAGTATTTGACCAACAGATGTGAGCAGCTCTA 738
DB 642 TCTGTGACAGCATGTGGTCTGCTGACTACAGAGCTTACACAGGCTTGTAAGTCCGACGA 701
QY 739 GGCAGCGGCTGTACCGGGTGGAGTGGCCCTACCTGGGCGGACGACATACAGAGGAGA 798
DB 702 GGCATTAATCTACTACCCAGGAGCCTATTCTCCTGCTGGTGGTATTTATGAGAGCCAT 761
QY 799 CTGAATCCGAGACAGATGCTCAATAGTGAAGGCTATGAGCGGAGCTGAGGCTCTGTG 858
DB 762 ATCACTGATGATCATTCCTCATCATGATGATGATGATGATGATGATGATGATGATGAT 821
QY 859 CGGC-----CTCCAGCGCGGAGCCTGGCGGCTCTCAAGAACAGAGAGCATGTGAGCAG 912
DB 822 CGGCCGAGCTCTCCAGCTCTCTTCAACCGACATACCTACTGACGTGAACGAGCAAGAAAG 881
QY 913 GGGATCCGCTGAGTGTGGAAGTGTGATGTCAGTGAAGTGTGAGTGTGAGTGTGAGTGTG 972
DB 882 CTAAATTAACCAATTAAGGAGATGATGACAGAGAGATTTGAGATATTTATTCATCC 941
QY 973 AAAGAGATCCGCGAGGCTTGGAGCTGCGCTGGGCTCCCTCCAGAGATACCTGTGAC 1032
DB 942 AAAGAGATTAAGAACAGATGTTGAATGCAATGTTGTGCACTTGGGGATTTCAAGAGAA 1001
QY 1033 TTATATGACAAACAGATGCTGCTGTGTGAGCAGAGCGGAGCCGAGCTCCGCACTTTC 1092
DB 1002 TTATATGACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1061
QY 1093 CCCACCTACTGCTGGGCTGAGAGTGAAGCAGCAAACTGTGAGAGAGCTGCAAGAGAAAC 1152

DB 1062 GAGCATGTTCTTGGGCTGCAATGGAATGCTCTCAACTTAGAGGACTTACAGAACGA 1121
QY 1153 AGGCTACCCACATCTTGAACATGCCCCGAGATTGACAACTTCTACCTGAGCGCTTC 1212
DB 1122 GGGTACCGTATATCTTGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 1181
QY 1213 ACCTACCAATGATGGGCTGCTGGATGAGAGTGGGCGGCGGAGTGGTGGTGGTGGTGG 1272
DB 1182 GAGTATCATTAATTCATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1241
QY 1273 GAGAGCAGCCCTTCATTTAGGCTGCAAGAGCAGAGGCGGAGGCGGAGGCGGAGGCGGAG 1332
DB 1242 GAGACTTACAAATTCATCTTAAAGCAAGAAACATGATGATTAATGCTTGTGCTGCTG 1301
QY 1333 AAGATGGCGCTCAGCGCTCAGCGGCGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1392
DB 1302 AAAATGGGGGTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1361
QY 1393 TGCAGCTGAGAGAGGCGCTGCGGAGGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAG 1452
DB 1362 TGAATCTGAGACCGAGCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1421
QY 1453 CTTGCTTCTGCGGCGGAGCTGAGATCTTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1512
DB 1422 CCAAGCTTCATGAGAGCACTGGAAGATATCAGGGGATCTTGTGCTGCTGCTGCTGCTG 1481
QY 1513 TGTGCGGAGAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1555
DB 1482 TTGATTCATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1524

RESULT 9
US-09-969-708-398/c
: Sequence 398, Application US/09969708
: Patent No. US20020102532A1
: GENERAL INFORMATION:
: APPLICANT: Augustus, Meena
: TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Sign
: TITLE OF INVENTION: Sets
: FILE REFERENCE: 689290-70
: CURRENT APPLICATION NUMBER: US/09/969.708
: CURRENT FILING DATE: 2001-10-03
: PRIOR APPLICATION NUMBER: US/60/237,606
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: US/60/237,608
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: US/60/237,425
: PRIOR FILING DATE: 2000-10-03
: NUMBER OF SEQ ID NOS: 658
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 398
: LENGTH: 334
: TYPE: DNA
: ORGANISM: Homosapiens
US-09-969-708-398

Query Match 11.6%; Score 313.2; DB 10; Length 334;
Best Local Similarity 96.1%; Pred. No. 2.8e-75;
Matches 321; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 2370 TGCATCTCTGCT 2429
DB 334 TGCATCTCTGCT 275
QY 2430 CCCCAAGGAGGAGGATCTCTAGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAT 2489
DB 274 CCCCAAGGAGGAGGAGGATCTCTAGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAT 215
QY 2490 GCGTTCCTTCATCT 2549
DB 214 CGCTTCCTTCATCTG 155
QY 2550 ACACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2609

```
Db 154 ACACACCCAGATACAGGCGACAGAGATAGTCTCTTTTGTCTGCGC 95
      |||
Oy 2610 CTCGGCCAGTACAGTCTTTCATAGCCTTACAGTATCTGGCTTTGACGAGAAATTAAC 2669
      |||
Db 94 CTCGGCAGTACAGTCTTTCATAGCCTTACAGTATCTGGCTTTGACGAGAAATTAAC 35
      |||
Oy 2670 ACATTTTCATAAAAA 2703
      |||
Db 34 ACATTTTCATTTTGGTATTTCATTTTCAAAAAATTA 1

RESULT 10
US-09-775-925-7
: Sequence 7, Application US/0975925
: Patent No. US20010049358A1
: GENERAL INFORMATION:
: APPLICANT: Lucbe, Ralf M.
: APPLICANT: Wei, Bo
: TITLE OF INVENTION: DSP-12 AND DSP-13 DUAL-SPECIFICITY
: FILE REFERENCE: 200125.420
: CURRENT APPLICATION NUMBER: US/09/775.925
: NUMBER OF SEQ ID NOS: 33
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 7
: LENGTH: 1052
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-775-925-7

Query Match 7.7%: Score 208.6; DB 10; Length 1052;
Best Local Similarity 56.7%: Pred. No. 1.1e-46;
Matches 385; Conservative 0; Mismatches 294; Indels 0; Gaps 0;

Oy 877 CCGGCGGCTCTCAGAACAGAGAGATGAGCAGCGCGATCCGTGCTGAGCTTGAAA 936
      |||
Db 187 CTTCTCAGACCTACTGAAAGTGAACGAAAGGCTAATTAACCAATTAAGGAG 246
      |||
Oy 937 GTGTGATGTAGTACCTGAGAGTGTACCTGCAAGAGATCCGCGAGCTCTGGAG 996
      |||
Db 247 ATCATGATGACAGAGATTTGGAGATATTACATCCAAAGATAGAACACAGATTGGAA 306
      |||
Oy 997 CTGCGCTGGGCTCCCTCCAGCAGTACCGTACTTCATGACAAACAGATGCTGCTG 1056
      |||
Db 307 ATGCAATGCTGTGACACTGCGGGAATTCAGGAATTTATGACATTAATGATAGTG 366
      |||
Oy 1057 CTGTGACAGCAGCGGAGCGGAGCTCCCGCATCTTCCCGACCTTACCTGGGCTCAGAG 1116
      |||
Db 367 ATCCTTGGTCAATGATAGCCCTACACAGATATTGACATGTGCTCTGGGCTCAGAA 426
      |||
Oy 1117 TGGACGCAACCTTGAGAGAGTGCAGAGAACAGGCTACCCACATCTTGAACAG 1176
      |||
Db 427 TGGAAATGCTTCAACTTGAAGAGCTTACAGAACGAGGGGTATATATTTGAATGTC 486
      |||
Oy 1177 GCCCGGAGATGACAACTTACCTGAGCGCTTACCTACCTACCAATTTGGGCTCTGG 1236
      |||
Db 487 ACTCAGAGATAGTAACTTCTTCCAGAGAGCTTTGAGTATACATTAACATTTGGGATAT 546
      |||
Oy 1237 GATGAGAGTGGCCAGCTGCTGCCGACCTGGAAGAGAGCAGCCGCTTCATTTGAGGCT 1296
      |||
Db 547 GATGAAGAGCAAGGATCTCTGGGCTGATGGAATGACACTTAATTAATCTCTCTAAA 606
      |||
Oy 1297 GGAAGAGCAGGCAACCCAGCTGCTGCTCCTGCACTGAAGATGGGCTCAGCCGCTCAGG 1356
      |||
Db 607 GGAAGAAACATGATTAATATGCTTGTGCACTGCAAAATGGGGTGAGCTCAGCC 666
      |||
Oy 1357 GCCAGAGTCTGGCTATGCTCATGAGAGTGAAGTGAAGTGGAGAGAGCCCTGCGC 1416
      |||
Db 667 TCCACCGTATGCTCATGAGAGTGAAGATATGCTGGAATCTGAGACGAGGCTATGAGC 726
      |||
Oy 1417 CACGTGACAGAGCTCGGCGCATCGCCCGCAACCTTGCTTCTGCGCAGCTGACAG 1476
      |||
```

```
Db 727 TATGGAAGAAAGACGAAAGCTTAACCAAGCCCAACCTCATGAGACACTGGAA 786
      |||
Oy 1477 ATCTACAGGAGATCTCTACAGGCGCAGAACCTTAGAGGTGTGGGAGGAGAAAGTGTAG 1536
      |||
Db 787 GAGTATCAGGGGATCTTCTGCTGCAAGCTTCTTAGGCTGATTCTAGTGAAGGAGCAAG 846
      |||
Oy 1537 CATGAGAGAGCAGCGCA 1555
      |||
Db 847 CCCTGGGAGAGAAAGCA 865

RESULT 11
US-09-808-701-12
: Sequence 12, Application US/09808701
: Patent No. US20020146757A1
: GENERAL INFORMATION:
: APPLICANT: Tang, Y. Tom
: APPLICANT: Zhou, Ping
: APPLICANT: Goodrich, Ryle
: APPLICANT: Asundi, Vinod
: APPLICANT: Zhao, Qing A.
: APPLICANT: Wang, Jian-Rui
: APPLICANT: Xue, Aidong J.
: APPLICANT: Ren, Feiyen
: APPLICANT: Wang, Dunrui
: APPLICANT: Chen, Rui-hong
: APPLICANT: Drmanac, Radoje T.
: TITLE OF INVENTION: No. US20020146757A1el Nucleic Acids and
: FILE REFERENCE: 790CIP2D
: CURRENT APPLICATION NUMBER: US/09/808.701
: PRIOR FILING DATE: 2001-03-14
: PRIOR APPLICATION NUMBER: 09/649,167
: PRIOR FILING DATE: 2000-08-23
: PRIOR APPLICATION NUMBER: 09/540,217
: NUMBER OF SEQ ID NOS: 17
: SOFTWARE: PL_FL_genes Version 2.0
: SEQ ID NO 12
: LENGTH: 969
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (34)..(765)
US-09-808-701-12

Query Match 6.0%: Score 163.4; DB 10; Length 969;
Best Local Similarity 58.5%: Pred. No. 2e-34;
Matches 306; Conservative 0; Mismatches 211; Indels 6; Gaps 1;

Oy 355 CCCGAGAACAGAGAGAGAGAGAGCAGCAGCTGCAACCTCATGTATACAGTGTGAGCGCG 414
      |||
Db 235 CACAAGACATGAGAGTATGCTTCAACATCTTCAAGTATATCAACTTCTGCTGTC 294
      |||
Oy 415 CAGAGTACATCGGCTGAGCAGCCAGCTGAGAGGACCCCGGCTTCCCGGCTCGCTAC 474
      |||
Db 295 GAAAGACAGATCAGTGTGACATGCGCCCTGGAAGAGCGCTGGGGGAGACCGGGTCCGCTAC 354
      |||
Oy 475 CTGCTGTGATTTCTACACGAGAGAGAGAGTCTGAGGCGAGAGTGAAGAGCGTCTCTCTG 534
      |||
Db 355 ATGTGTGTGTGTACAGCAGCGGCGCCAGGA-----CACGAGAGAAATATCTTGTG 408
      |||
Oy 535 GCGGTGATTTCCCTGACAGCAGCTCCCGCAGCTGACCTGGGCTGTGGTGGCCCTC 594
      |||
Db 409 GAGGTGACTTTTCCAGTAAAGAACTAAAGCTGACCATTTGGAGTGTCTCGCACTG 468
      |||
Oy 595 TGGAGTACACCCAGGCTGACTTATGAGAGAGGAGGCTTGAAGCTGAGCTGTGAGTGG 654
      |||
Db 469 TGGAGGACACGAAATTCACCTGTATGAGAGATGTTGGGTTTCAAGCTGAGACAGCAGGA 528
      |||
Oy 655 CAAAGCGGATCTTCAAGCCCATCTTCATTCAGAACCATATGTGGGCGCACTCCAGGATTTG 714
      |||
```

D _b	529	AGGATGCACATATTAAAGCCGTGTCTGTCCAGGCCATGTGGTCTGCCCTCGACAGTGCTT	588
O _y	715	CACCAGCATGTGAGGCAGCTCTTAGCAGCGGCTTTACCGSGTGGCAGTGCCCTCAC	774
D _b	589	CACAAGGCTCTCGAAGTGGCCCCGGAGGCACAACTACTTCCC GGSGSTGTAGTCGTATC	648
O _y	775	TGGGCCACGCCACTACCGAGAGACTGAACCTCCGACAGAGCTGCCCAATGAGTGSACG	834
D _b	649	TGGGCTACTACTTGTGAGAGCTGATCAGCTCCGAGCAGAGCTGCATCAACGAGTGS AAC	708
O _y	835	GCTATGGCCGACTGGAGTCTCTCGGGCTTCCACGCGCGAGC	877
D _b	709	GCCATGCAAGAGCTGGAGTCTACGCGGGCCGACCTCCCCCGGC	751

```

RESULT 12
US-10-002-600-51
, Sequence 51, Application US/10002600
, Patent No. US20020137077A1
, GENERAL INFORMATION:
, APPLICANT: Hopkins, Christopher M.
, APPLICANT: Peterson, David P.
, APPLICANT: Cocks, Benjamin G.
, APPLICANT: Hawkins, Phillip R.
, TITLE OF INVENTION: GENES REGULATED IN ACTIVATED T CELLS
, FILE REFERENCE: PA-0042 US
, CURRENT APPLICATION NUMBER: US/10/002,600
, CURRENT FILING DATE: 2001-10-25
, PRIOR APPLICATION NUMBER: 60/243,521
, PRIOR FILING DATE: 2000-10-25
, NUMBER OF SEQ ID NOS: 116
, SOFTWARE: PERL Program
, SEQ ID NO 51
, LENGTH: 1685
, TYPE: DNA
, ORGANISM: Homo sapiens
, FEATURE:
, NAME/KEY: misc.feature
, OTHER INFORMATION: Template ID: 235631.2
US-10-002-600-51

```

Query Match	3.7%;	Score 98.8;	DB 12;	Length 1685;
Best Local Similarity	53.3%;	Pred. No. 8.4e-17;		
Matches 208;	Conservative	0;	Mismatches 182;	Indels 0;
				Gaps 0.

Qy	1087	ATTTTCCCCCCTCTACCTCGGGCTAGAGTGGAGACGCAACACTGAGAGCGTCGAG	1146
Db	613	ATCTTGCCCTACTACCTTTCCTGGGAGCTGCAGTCACTCTGTCAGACCTGCAAGGGGCTGAC	672
Qy	1147	AGGAACAGGAGTACCACATCTTGAACATGGCCCCGGAGATATGACAACTTTCACCTGAG	1206
Db	673	GGCTGTGGCATTACAGGCCGTCTCAACGTGTCCGACGTGCCCCACCACTTTTGAAGGC	732
Qy	1207	CGCTTCACCTACACCAATGTGCGCTCTGGGATGAGAGATCGGCCACGTCGTGGCCGAC	1266
Db	733	CTTTTCCCTCTACAAGAGATTCCTCTGTGAGAGAACACAGATGGTGAATCATGTCCTGG	792
Qy	1267	TGGAGGAGAGACGACCGCTCATTTGAGGCTGTGAAGAGACAGAGGACCCACGACTGCTGTC	1326
Db	793	TTTCAGGAGGCCATTAGGCTTATTGACTGTGGGTAAAGAACAGCGGAGGCGGGTCTGTGTG	852
Qy	1327	CACCTCAGATGTGGCGTTCAGCCGCTCAGCGGCCACCAATGCTGGCCCTTATGCATGAAGAG	1386
Db	853	CACCTCCAGGGGGGATTCGTGCGCTGTGCCACCATCTGTCTGGCATTACTCATGCAAGAT	912
Qy	1387	TACGATGACGCTGTGAGACAGGCCCCCTGCGCACACTGACAGAGTCTCGGCCCCATTCGCCCG	1446
Db	913	CGCCCTGTGCGGCTGAGAGAGCCCTTTGACTTGTTAAGCAGCGCCGGGGGGGTCAATCTCC	972
Qy	1447	CCCAACCGCTGGTCTCTGCGGCACCTCCAG	1476
Db	973	CCCAACTTCAGTTTCATGGGCACTGCTG	1002

```

RESULT 13
US-09-865-993-3
: Sequence 3, Application US/09865993
: Publication NO. US20030060437A1
: GENERAL INFORMATION:
: APPLICANT: Brett P. Monia
: APPLICANT: Andrew T. Walt
: TITLE OF INVENTION: ANTISENSE MODULATION OF DUAL SPECIFIC PHOSPHATASE 5 EXPRESSION
: FILE REFERENCE: RTS-0175
: CURRENT APPLICATION NUMBER: US/09/865,993
: CURRENT FILING DATE: 2001-05-25
: NUMBER OF SEQ ID NOS: 89
: SEQ ID NO 3
: LENGTH: 2139
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (47)..(1201)
US-09-865-993-3

```

Query Match	3.3%	Score 89.2;	DB 9;	Length 2139;
Best Local Similarity	51.5%;	Pred. No. 3.7e-14;		
Matches 205; Conservative	0;	Mismatches 193;	Indels 0;	Gaps 0;

Oy	1087	ATTTTCCCCACCTTCACCTCGGGGCTCGAGATGAGAACGACCAACCTGAGAGAGCTGCAG	1146
Db	950	ATTCCTTCCCTTCTCTCTCTCTTTGGAGTGCTTACCATGCACTCCCAAGTGGAGTTCCTCCGC	649
Oy	1147	AGGAACAGGGTCACCCCACTCTTGAACATGGGCCGGGAGATTGACAACCTTCAACCTGAG	1206
Db	650	AACCTTGCACATACAGGCCCTGCTGAATGTCCTCCGACGGACCTCCGAGAGCCTGCATGACC	709
Oy	1207	CGCTTCACCTACCAACAAATGTCGGCTCTGGGATGAGAGATGGGCCAGCTGCTGGCGAC	1266
Db	710	CACCTACACTCAAAATGATGCCCTGTGGAGACAGCCACAGCGCTGCATTAAGTCCACC	769
Oy	1267	TGGAAGGAGACGCAACCGCTCATTTGAGAGGCTGCAAGAGACAGGCAACCCAGCTGCTGTGTC	1326
Db	770	TTTCAAGAGCAATAGACTTCCTATTTGACTGTGTGTAAGGAAAAGGAGGACAGGTCTCTGGTCC	829
Oy	1327	CACCTCAGATGAGGGCGTCAGCCGCTCAGCGGCCACAGTGTGGCTTATGCAATGCAAGCAG	1386
Db	830	CACCTGTGAGGCTGGGGATCTCCCGTTACACCACCATCTGCATGGCTTACTTATTAAGAAGCC	889
Oy	1387	TACGATCAGCAGCTTGAGAGCAGGCGCTTGGCCGACAGCTGAGAGAGCTCCGGCCCATTCGCCGC	1446
Db	890	AAGCGATTCCCTCGAAGAGAGCGCCTTGATTAATCAATCAACAGAGGAGGAGCATGGTCTTCG	949
Oy	1447	CCCAACCGCTTCCTCGCGGCACTCTCAGATTCATACCA	1484
Db	950	CCCACTTTGGCTTCATATGAGGCACTCTCTCAGATACGA	987

```

RESULT 14
US-09-865-993-10
; Sequence 10, Application US/09865993
; Publication No. US20030060437A1
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: Andrew T. Watt
; TITLE OF INVENTION: ANTISENSE MODULATION OF DUAL SPECIFIC PHOSPHATASE 5 EXPRESSION
; FILE REFERENCE: RTS-0175
; CURRENT APPLICATION NUMBER: US/09/865,993
; CURRENT FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 89
; SEQ ID NO 10
; LENGTH: 2479
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

```

US-09-865-993-10

Query Match	3.3%	Score 89.2;	DB 9;	Length 2479;
Best Local Similarity	51.5%;	Pred. No. 4e-14;		
Matches 205; Conservative	0;	Mismatches 193;	Indels 0;	Gaps 0;

QY	1087	ATTTCCTCCACCTCTACCTGGGCTCAGAGTGGAAAGCAGCAAACTGTGGAGAGCTGCAG	1148
Db	758	ATCTCTCCCTCTCTCTACTTACCTTGGAAAGTGGCTACATGATCCAAATGGGAGTCTCTCGCC	817
QY	1147	AGGAACAGGGGTACCCACACTCTTGAACATGGCCCGGAGATTGACAACTTCTACCCCTGAG	1206
Db	818	AACGTGCACATACAGACCCCTGGCTGAAATGTCTCCGACGGAGCTCGAGAGCTGCATGAC	877
QY	1207	CGCTTACACTACACAAATGTGGCCTCTGTGGATGAGAGTGGGCCAGCTGCTGCTGCAC	1266
Db	878	CACCTACCTACAAATGAGATCCCTGTGGAAGACAGCAACCGCTGCATTAAGCTCCAC	937
QY	1267	TGGAAAGAGACGACCGCTCTATTGTAGGGCTGGCAAGACACAGGACACCAGCTGCTGTGCT	1326
Db	938	TTTTCAGAACCAATAGACTTTCATTGTACTGTGTGTGAGGAAAAGGAGGCAAGTCTCTGTGCT	997
QY	1327	CACGTCAAGATGGGGCTCTACGCCCTCATGGCGCACAGTGTCTGGCTTATGCCATGAAGAG	1386
Db	998	CACGTGAGGGCTTGGATCTCCGCTTACCCACCACATCTGCATGTGGCTTACTTATGAAGAC	1057
QY	1387	TACGAATGACCTGTGAGCAGGCCCTGGCGCAAGTCGACGAGAGCTCGGCGCATCGCGGC	1446
Db	1058	AACGAGTTCGCGCTGAGGAGGCGCTTCGATTACATCAAGACAGGAGGAGCATGGTCTCG	1117
QY	1447	CCCAACCTGGCTTCTCGCGCCAGCTGCAGATCTACCA	1484
Db	1118	CCCAACTTGTGGCTTCATGGGCGACGCTCTCGACATGAGA	1155

RESULT 15
US-09-902-941-1918

```

APPLICANT: Henderson, Robert A.
APPLICANT: Wang, Tongtong
APPLICANT: Watanabe, Yoshihiro
APPLICANT: Johnson, Jeffrey C.
APPLICANT: Keltner, Marc W.
APPLICANT: Mainerakis, Margarita
APPLICANT: Carter, Darrick
APPLICANT: Fanger, Gary R.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Bangur, Chantanya S.
APPLICANT: McNabb, Andia
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.478C17
CURRENT APPLICATION NUMBER: US/09/902,941
CURRENT FILING DATE: 2001-07-10
NUMBER OF SEQ. ID NOS.: 2002
SOFTWARE: FastSeq for Windows Version 4.0
SEQ. ID NO. 1918
LENGTH: 1209
TYPE: DNA
ORGANISM: Homo sapiens
IS-09-902-941-1918

```

Query Match	3.3%	Score 88;	DB 9;	Length 1209;
Best Local Similarity	50.5%;	Pred. No. 6.3e-14;		
Matches 214; Conservative	0;	Mismatches 210;	Indels 0;	Gaps 0;

QY 1087 ATTTCCTCCACCTTACCTTACCTGGGCTAGAGTGGAAACGGACGCAAACTGGAGAGCTGGAC 1146

Db 616 ATCTCTCCCTTCTTACTCTGCGAGTGGCTTACCACTGCTGCTCCGGAGMACATGTTGTGGAC 675

Y 1147 AGGAACAGGGTCCACACATCTTTGAACATGAGCCGGAGGAGATTGACAACTTCAACCTTAG 1208

Db	676	GCCTTGGGATCAAGGCTCTGTGAATGTCTCTCGAGATGTGCCCAACACATTGTAAGA	735
QY	1207	CGCTTACACCTTACCAACAATGTGCCCTCTGTGGATGAGGAAGTGGCCAGCTGCTGCCGAC	1266
Db	736	CACATTCAGTACAAGATGCATCCAGTGAAGATTAACCAACAAAGGCCGACATCAGCTCTGG	795
QY	1267	TGAGAGAGAGACGACCGCTCTATTGAGGCTGCACAAAGACACAGGGCACCCACGCTGTGTC	1326
Db	796	TTCTATGTAAGACCATATAGAGTACATCGATGCCGTGTAAAGACTGCGCGTGGGCGCTGTGGTG	855
QY	1327	CACATGCAAGATGGGCGTCAAGCCGCTCAGCGGCCACAGTGTGGCTATGTCATGAAGACAG	1386
Db	856	CATCGCCAGCGGGCATCTCGCGGTGGGCCACCATCTGCGCTGACCTGATCATGTAAAG	915
QY	1387	TACGAATGCAAGCTGTGAGACGAGGGCCCTGCGGCACAGTSCAAGAGACTCCGGCCATCGCCCGC	1446
Db	916	AAACGGGTGAGGCTGTGGAGGGAGGCGCTTCGAGTTCTGTTAAACACACGCGCAGCATCTCATCG	975
QY	1447	CCCAACCCCTGGCTTCTCGAGCCCAAGCTCAGAGATCTTACACGAGGATCCTTGACGCCGAAC	1506
Db	976	CCCAACCTTCAGCTTCATGAGGGGACAGCTGTGTGCAAGTTGAGATCCAGAGTGTGTGCCACATCG	1035
QY	1507	TGAG 1510	
Db	1036	TGTG 1039	

Search completed: April 12, 2003, 07:29:43
Job time : 426.978 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 10, 2003, 23:58:55 ; Search time 7346.79 Seconds

(without alignments)
5960.776 Million cell updates/sec

Title: US-09-761-640-1

Perfect score: 2704
Sequence: 1 cgtccctcctgctcgtcg...aaaaaaaaaaaaaaaaa 2704

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapect 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :
EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estlmu:*
5: em_estlov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vit:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	839	31.0	991	13	BM471207 AGENCOURT
2	767.2	28.4	875	12	BG752969 602732427
3	758	28.0	812	12	BG750408 602709243
4	755.2	27.9	790	12	BG326709 60242578
5	741.6	27.4	772	14	BQ014446 UT-H-ED1-
6	727.8	26.9	1131	12	BE737389 601306112

7	698.6	25.8	846	10	BE563259
8	695.6	25.7	932	12	BG340220
9	673.8	24.9	726	12	BE869222
10	664	24.6	910	12	BG169297
11	662	24.5	663	12	BG753072
12	652.4	24.1	1004	14	BM910493
13	642	23.7	653	12	BG679981
14	641.6	23.7	919	12	BE737087
15	635.4	23.5	716	12	BE907514
16	622	23.0	890	12	BE901875
17	619.2	22.9	897	12	BE873337
18	609.4	22.5	672	12	BG831015
19	600.4	22.2	1228	14	BM921421
20	599.4	22.2	1125	14	BM809903
21	599	22.2	599	10	BE280203
22	598.4	22.1	678	13	BM047611
23	596	22.0	1024	13	BM560192
24	590.4	21.8	622	12	BE898514
25	586.4	21.7	589	9	AL602806
26	577.6	21.4	1061	14	BM906413
27	567.2	21.0	728	13	B1824278
28	562.6	20.8	791	12	BE881145
29	561.8	20.8	892	12	BE971565
30	548.8	20.3	917	14	BQ712114
31	548.4	20.3	584	13	BM018531
32	541.6	20.0	863	12	BG749236
33	537.6	19.9	969	10	BE563116
34	535.4	19.8	943	13	BM563401
35	517.4	19.1	574	13	B1007055
36	517	19.1	937	14	BQ718968
37	516	19.1	719	10	BE543087
38	505	18.7	750	10	BE531347
39	502	18.6	846	13	B1820326
40	490.2	18.1	560	13	B1007133
41	482.4	17.8	568	13	B1087083
42	481.6	17.8	496	10	BE348287
43	471	17.4	505	13	BM083928
44	469.8	17.4	952	14	BQ675874
45	468	17.3	729	13	B1086720

ALIGNMENTS

RESULT 1
BM471207
LOCUS
DEFINITION
AGENCOURT.6478155 NIH_MGC_72 Homo sapiens CDNA clone IMAGE:5563111
5', mRNA sequence.
BM471207
BM471207.1 GI:18520249
EST.
SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 991)
NIH-MGC <http://mgc.ncl.nih.gov/>.
AUTHORS
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC/DCTP/DTF
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMNL at:
<http://image.llnl.gov>
Plate: LMNL2293 row: 3 column: 08
High quality sequence stop: 592.
Location/Qualifiers
1. .991

FEATURES
source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:556311"
/clone_lib="NIH_MGC_72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."

```

```

BASE COUNT      199 a      303 c      329 g      159 t      1 others
ORIGIN

```

```

Query Match      31.0%; Score 839; DB 13; Length 991;
Best Local Similarity 93.5%; Pred. No. 1,1e-116;
Matches 920; Conservative 0; Mismatches 56; Indels 8; Gaps 4;

```

```

QY 40 GTTGGAGGAAGGGCCGTCGACCCGACGCGCTGCTCCAGAGCTTGGCTGCTCCATGGCC 99
DB 1 GTTGGAGGAAGGGCCGTCGACCCGACGCGCTGCTCCAGAGCTTGGCTGCTCCATGGCC 60
QY 100 CTGCTCACTAGTACGCTTGGCCCGGCGAGCGGCTCTCCAGCGCCCTGGGGCCCTG 159
DB 61 CTGCTCACTAGTACGCTTGGCCCGGCGAGCGGCTCTCCAGCGCCCTGGGGCCCTG 120
QY 160 GACGAGCGGCTCCAGGAGAGTCCAGCTCCAGGAGAGAGCTTGGCTGCTCCATGGCC 219
DB 121 GACGAGCGGCTCCAGGAGAGTCCAGCTCCAGGAGAGAGCTTGGCTGCTCCATGGCC 180
QY 220 GGGGCTGCTCTGGGACTGAGATGAGGAGGAGCAATGATGATGATGATGATGATGATGAT 279
DB 181 GGGGCTGCTCTGGGACTGAGATGAGGAGGAGCAATGATGATGATGATGATGATGATGATGAT 240
QY 280 GAGCCAAAGAGAGAGGCGCCGAGTGGAGAGAGCTCCAGGGGAGACAGACATTCGCG 339
DB 241 GAGCCAAAGAGAGAGGCGCCGAGTGGAGAGAGCTCCAGGGGAGACAGACATTCGCG 300
QY 340 CAAGGATCCCAAGTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 399
DB 301 CAAGGATCCCAAGTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
QY 400 CAGCTCTAGAGCGCGAGATGACATCCGCTTGGCAGCCGAGCTGAGAGCACCAGCGCT 459
DB 361 CAGCTCTAGAGCGCGAGATGACATCCGCTTGGCAGCCGAGCTGAGAGCACCAGCGCT 420
QY 460 CCGCGGCTCCGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 519
DB 421 CCGCGGCTCCGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 520 GAGAGGAGCTCTGAGAGTGGATTTCCCTGACAGCAGCTCCCGAGCTGACACCTGGGC 579
DB 481 GAGAGGAGCTCTGAGAGTGGATTTCCCTGACAGCAGCTCCCGAGCTGACACCTGGGC 540
QY 580 CTGCTCTTGGCCCTCTGAGAGTGGATTTCCCTGACAGCAGCTCCCGAGCTGACACCTGGGC 639
DB 541 CTGCTCTTGGCCCTCTGAGAGTGGATTTCCCTGACAGCAGCTCCCGAGCTGACACCTGGGC 600
QY 640 GTGACGCTTGGTGGGCAAGCCGAGATCTTCAAGCCATCTCCATCCAGACATCTGGGCC 699
DB 601 GTGACGCTTGGTGGGCAAGCCGAGATCTTCAAGCCATCTCCATCCAGACATCTGGGCC 660
QY 700 ACACCTCAGGATTTGACACAGCATGAGGAGAGCTTAGGACAGCGCTTGTACCGGG 758
DB 661 ACACCTCAGGATTTGACACAGCATGAGGAGAGCTTAGGACAGCGCTTGTACCGGG 720
QY 759 TGGGAGTGGCCCTACCTGGGCGACCACTACAGAGAGAGTGAATCTCCAGACAGAGCTG 818
DB 721 TGGGAGTGGCCCTACCTGGGCGACCACTACAGAGAGAGTGAATCTCCAGACAGAGCTG 780
QY 819 CTTCAATGAGTGGAGGCTATGGCCGAGCTGAGAGTCTTGGCTCCAGCGCCGAGGCC 878
DB 781 CTTCAATGAGTGGAGGCTATGGCCGAGCTGAGAGTCTTGGCTCCAGCGCCGAGGCC 840

```

```

QY 879 TGGGAGTGGCTCTAG-AAACAGAGACAGATGAGAGGAGGATCGGTGAGC---TGTGGA 934
DB 841 TGGGAGTGGCTCTAG-AAACAGAGACAGATGAGAGGAGGAGGATCGGTGAGC---TGTGGA 900
QY 935 AAGTGTGAGATGTCAGTACCTGAGAGTGTCACTTCC--AAAGAGATCCGACAGCTC 991
DB 901 AAGTGTGAGATGTCAGTACCTGAGAGGAGGATCGGTCAAAAAGAGAGAGAGAGAGAGAG 960
QY 992 TGGAGCTGCGCTGAGGAGCTCCGCC 1015
DB 961 TGGAGCTGCGCTGAGGAGCTCCGCC 984

```

```

RESULT 2
BG752969
LOCUS
DEFINITION
602732427F1 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:4876014 5',
ACCESSION
BG752969
VERSION
BG752969.1 GI:14063622
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 (bases 1 to 875)
NIH-MGC http://mgi.nci.nih.gov/.
NATIONAL INSTITUTES OF HEALTH, MAMMALIAN GENE COLLECTION (MGC)
UNPUBLISHED (1999)
CONTACT: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LCM1758 row: e column: 07
High quality sequence stop: 818.
Location/Qualifiers
1..875

```

FEATURES

```

source
1..875

```

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4876014"
/clone_lib="NIH_MGC_43"
/tissue_type="normal pigmented retinal epithelium"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pOTR7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(5). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

```

```

BASE COUNT      184 a      267 c      275 g      149 t
ORIGIN

```

```

Query Match      28.4%; Score 767.2; DB 12; Length 875;
Best Local Similarity 97.0%; Pred. No. 6.7e-106;
Matches 846; Conservative 0; Mismatches 18; Indels 8; Gaps 6;

```

```

QY 289 GAGAAAGCCCCGAGTAGAGAGCTCCAGCGGAGACACAGACTTGGGCAAGATCC 348
DB 1 GAGAAAGCCCCGAGTAGAGAGAGCTCCAGCGGAGACACAGACTTGGGCAAGATCC 60
QY 349 CAGAGTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 408
DB 61 CAGAGTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
QY 409 AGCGCGAGAGATGACATCCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 468

```

```
Dh 121 AGCCGCGAGATGACATCCGCCCTGGCAGCCAGCTGAGGACACCCCGGCTCCGCCGCTC 180
Qy 469 CGTACCTGCTGTAGTATTCTACAGAGAGAGAGAGAGTCTGAGCCAGATGAGACGGTC 528
Dh 181 CGCTACCTGCTGTAGTATTCTACAGAGAGAGAGAGAGTCTGAGCCAGATGAGACGGTC 240
Qy 529 CTCCTGGGCGGTGATTTTCCCTGACAGACAGCTCCGCCAGCTCCAGCCCTGGCTGTG 588
Dh 241 CTCCTGGGCGGTGATTTTCCCTGACAGAGCTCCGCCAGCTCCAGCCCTGGCTGTG 300
Qy 589 CCCCTCTGAGATGACACCCAGAGTGTACTATGAGAGAGGGGGCTTACGGTACGCTCT 648
Dh 301 CCCCTCTGAGATGACACCCAGAGTGTACTATGAGAGAGGGGGCTTACGGTACGCTCT 360
Qy 649 GGTGGGCAAAAGCCGATCTTCAAGCCCATCTCATCCAGACATGTGGGCCACACTCCAG 708
Dh 361 GGTGGGCAAAAGCCGATCTTCAAGCCCATCTCATCCAGACATGTGGGCCACACTCCAG 420
Qy 709 GTATTGACCAAGCATGTGAGGAGCAGCTCTAGGACGGGCTTTGACCGGGTGGCAGTCC 768
Dh 421 GTATTGACCAAGCATGTGAGGAGCAGCTCTAGGACGGGCTTTGACCGGGTGGCAGTCC 480
Qy 769 CTCACCTGGGCGAGCACTTACAGAGAGAGTGAACCTCCGAACAGAGCTGGCTCAATAG 828
Dh 481 CTCACCTGGGCGAGCACTTACAGAGAGAGTGAACCTCCGAACAGAGCTGGCTCAATAG 540
Qy 829 TGAAGCGGTATGCGGAGCTGAGTCTCTGGGCGCTCCAGCGGCGAGCTGGCGGCTCC 888
Dh 541 TGAAGCGGTATGCGGAGCTGAGTCTCTGGGCGCTCCAGCGGCGAGCTGGCGGCTCC 600
Qy 889 TCAGAACAGAGAGAGATGAGAGAGCGAGTCGCTGAGTGTGAAAGTGTGGATGTC 948
Dh 601 TCAGAACAGAGAGATGAGAGAGCGAGTCGCTGAGTGTGAAAGTGTGGATGTC 660
Qy 949 AGTACCTGGA-GAGTGTCACTTCCAAAGAGATCCGCCAGGCTCTGGAGCTGGCGCTGG 1007
Dh 661 AGTACCTGGA-GAGTGTCACTTCCAAAGAGATCCGCCAGGCTCTGGAGCTGGCGCTGG 720
Qy 1008 GCTCCCGCC-TCAGAGAGTACGCTGACTCATGACAAAC---AGATGTGCTGCTGGTGG 1063
Dh 721 GCTCCCGCC-TCAGAGAGTACGCTGACTCATGACAAAC---AGATGTGCTGCTGGTGG 780
Qy 1064 CACAGCGGAGCCGAGCTCCGAGTCTTCCGCCACCTTACCTGGGCTCAGAGTGAAGC 1123
Dh 781 CACAGCGGAGCCGAGCTCCGAGTCTTCCGCCACCTTACCTGGGCTCAGAGTGAAGC 838
Qy 1124 CAGCAACCTGGAGAGCTGCAAGAGAACAGG 1155
Dh 839 CAG-AAACCTTGGAGAGCTGCCGAGGAACAGG 869

RESULT 3
Bg750408 812 bp mRNA linear EST 15-MAY-2001
LOCUS Bg750408 60270924F1 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:4845707 5',
DEFINITION mRNA sequence.
ACCESSION Bg750408
VERSION Bg750408.1 GI:14061061
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 812)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: rgs@bbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
```

```
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
plate: LNCM685 row: f column: 12
High quality sequence stop: 792.
FEATURES
    source
        1..812
            location/Qualifiers
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:4845707"
                /clone_lib="NIH_MGC_43"
                /tissue_type="normal pigmented retinal epithelium"
                /lab_host="DH10B (phage-resistant)"
                /note="Organ: eye; Vector: pOTB7; Site 1: XhoI; Site 2:
                EcoRI; cDNA made by oligo-dT priming. Directionally
                cloned into EcoRI/XhoI sites using the following 5'
                adaptor: GGCACGAG(G). Library constructed by Ling Hong
                in the laboratory of Gerald M. Rubin (University of
                California, Berkeley) using ZAP-cDNA synthesis kit
                (Stratagene) and Superscript II RT (Life Technologies).
                Note: this is a NIH_MGC library. |"
BASE COUNT 167 a 250 c 254 g 141 t
ORIGIN
Query Match 28.0%; Score 758; DB 12; Length 812;
Best Local Similarity 97.3%; Pred. No. 1.7e-104;
Matches 792; Conservative 0; Mismatches 20; Indels 2; Gaps 2:
Qy 289 GAGAAAGCCCCGAGTGAAGAGAGCTCCACGGGGACCAAGACAGATTGGGCAAGATCC 348
Dh 1 GAGAAAGCCCCGAGTGAAGAGAGCTCCACGGGGACCAAGACAGATTGGGCAAGATCC 60
Qy 349 CAGAGTCCCCGAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 408
Dh 61 CAGAGTCCCCGAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
Qy 409 AGGCGCGAGATGACATCCGCTGGCAGCCAGCTGAGAGAGACCCGGGCTCCCGGCTC 468
Dh 121 AGGCGCGAGATGACATCCGCTGGCAGCCAGCTGAGAGAGACCCGGGCTCCCGGCTC 180
Qy 469 CGCTACCTGTGTGATGTTTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 528
Dh 181 CGCTACCTGTGTGATGTTTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
Qy 529 CTCCTGGGCGGTGATTTTCCCTGACAGAGAGCTCCGCCAGCTCCAGCCCTGGCTGTG 588
Dh 241 CTCCTGGGCGGTGATTTTCCCTGACAGAGAGCTCCGCCAGCTCCAGCCCTGGCTGTG 300
Qy 589 CCCCTCTGAGATGACACCCAGAGTGTACTATGAGAGAGAGGGGGCTTACGGTACGCTCT 648
Dh 301 CCCCTCTGAGATGACACCCAGAGTGTACTATGAGAGAGAGGGGGCTTACGGTACGCTCT 360
Qy 649 GGTGGGCAAAAGCCGATCTTCAAGCCCATCTCATCCAGACATGTGGGCCACACTCCAG 708
Dh 361 GGTGGGCAAAAGCCGATCTTCAAGCCCATCTCATCCAGACATGTGGGCCACACTCCAG 420
Qy 709 GTATTGACCAAGCATGTGAGGAGCAGCTCTAGGACGGGCTTTGACCGGGTGGCAGTCC 768
Dh 421 GTATTGACCAAGCATGTGAGGAGCAGCTCTAGGACGGGCTTTGACCGGGTGGCAGTCC 480
Qy 769 CTCACCTGGGCGAGCACTTACAGAGAGAGTGAACCTCCGAACAGAGCTGGCTCAATAG 828
Dh 481 CTCACCTGGGCGAGCACTTACAGAGAGAGTGAACCTCCGAACAGAGCTGGCTCAATAG 540
Qy 829 TGAAGCGGTATGCGGAGCTGAGTCTCTGGGCGCTCCAGCGGCGAGCTGGCGGCTCC 888
Dh 541 TGAAGCGGTATGCGGAGCTGAGTCTCTGGGCGCTCCAGCGGCGAGCTGGCGGCTCC 600
Qy 889 TCAGAACAGAGAGAGATGAGAGAGCGAGTCGCTGAGTGTGAAAGTGTGGATGTC 948
Dh 601 TCAGAACAGAGAGATGAGAGAGCGAGTCGCTGAGTGTGAAAGTGTGGATGTC 660
Qy 949 AGTACCTGGAAGTGTCACTTCCAAAGAGATCCGCCAGGCTCTGGAGAGCTCGGCTGGGG 1008
```

|||||
Db 661 AGTGAAGTGGAGAGTGTCTCAATGAGATCCGACAGGCTGTGAGCTGTGCGCTGGG 720
QY 1009 CTCCTCCACACAGTACGCTGATCATCGACACCAATGCTGCTGGTGGACAG 1068
|||||
Db 721 TGTCCCTTCACAGTACGCTGATCATCGACAA-CAGATGCTGCTGTGGTGGACAG 779
QY 1069 CGGAGCCGAGCTCCGCTGCTTCCCGCCACTCT 1102
|||||
Db 780 C-GGAACGAGCTCCCGATTTTCCCACTCTACT 812

RESULT 4
BG326709 790 bp mRNA linear EST 27-FEB-2001
LOCUS 603425578F1 NIH_MGC_14 Homo sapiens cDNA clone IMAGE:4563595 5',
DEFINITION mRNA sequence.
ACCESSION BG326709
VERSION BG326709.1 GI:13133146
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 790)
NIH-MGC http://mgc.ncl.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DRP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L10M1277 row: k column: 20
High quality sequence stop: 737.
Location/Qualifiers
1..790
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4563595"
/clone_lib="NIH_MGC_14"
/tissue_type="renal cell adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: kidney; Vector: pOT7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 162 a 237 c 268 g 123 t
ORIGIN
Query Match 27.9%; Score 755.2; DB 12; Length 790;
Best Local Similarity 98.2%; Pred. No. 4.4e-104;
Matches 774; Conservative 0; Mismatches 13; Indels 1; Gaps 1;

QY 45 GGAAGAGGCGGTGCGGTGCGACCCAGAGTGTGCGGCTGCTCCATGGCTGTG 104
|||||
Db 1 GGAAGAGGCGGTGCGGTGCGACCCAGAGTGTGCGGCTGCTCCATGGCTGTG 60
QY 105 CACAATGAGCGTTGCCCCGGGAGGCGGCTTCACAGCCCTGGGGCCTGGGACCA 164
|||||
Db 61 CACAATGAGCGTTGCCCCGGGAGGCGGCTTCACAGCCCTGGGGCCTGGGACCA 120
QY 165 GCGGCTCAGAGGAGGATGCAGCTCCAGAGGAGGAGAGCTTGGGGTCTCCGTGGGCG 224
|||||
Db 121 GCGGCTCAGAGGAGGATGCAGCTCCAGAGGAGGAGAGCTTGGGGTCTCCGTGGGCG 180

QY 225 TGNCTGGGAGTGGAGATGGAGGAGCAATGATGATGACAGCAGAGCCAGTTCTAGCC 284
|||||
Db 181 TGTCTGGGAGCTGGCAGATGGAGGAGCAATGATGATGACAGGAGCCAGTTCTAGCC 240
QY 285 AACAGAGAGGCCCCAGTGGAGGAGAGGCTCCACGGGAGCACAGACAGACTTGGGCAAG 344
|||||
Db 241 AACAGAGAGGCCCCAGTGGAGGAGAGAGCTCCACGGGAGCACAGACAGACTTGGGCAAG 300
QY 345 ATCCAGAGTCCCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 404
|||||
Db 301 ATCCAGAGTCCCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
QY 405 GCTGAGGCGCCAGAGATGACATCCGCTGGCAGGCCAGCTGGAGGAGGAGGAGGAGGAGG 464
|||||
Db 361 GCTGAGGCGCCAGAGATGACATCCGCTGGCAGGCCAGCTGGAGGAGGAGGAGGAGGAGG 420
QY 465 GCTCCGCTACCTGCTGCTGATGATTTCTACAGAGAGAGAGGAGGAGGAGGAGGAGGAGG 524
|||||
Db 421 GCTCCGCTACCTGCTGCTGATGATTTCTACAGAGAGAGAGAGGAGGAGGAGGAGGAGG 480
QY 525 GGTCTCTCTGGGCGTGGATTTCTGACAGAGAGCTCCCGAGCTGACCCCTGGCTGGT 584
|||||
Db 481 GGTCTCTCTGGGCGTGGATTTCTGACAGAGAGCTCCCGAGCTGACCCCTGGCTGGT 540
QY 585 CTTCGCTCTGAGAGTACACAGCCAGGCTGATGATGAGAGGAGGAGGAGGAGGAGGAGG 644
|||||
Db 541 CTTCGCTCTGAGAGTACACAGCCAGGCTGATGATGAGAGGAGGAGGAGGAGGAGGAGG 600
QY 645 GTCTGTGGGCAAGGCGGATCTTCAAGCCCATCTCCATCCAGAGACCATGTGGGCGCAC 704
|||||
Db 601 GTCTGTGGGCAAGGCGGATCTTCAAGCCCATCTCCATCCAGAGACCATGTGGGCGCAC 660
QY 705 CCAGGATTTGACAC -AAGCATGTGAGGAGGCTGTAGGAGCGGCTTGTACCGGGTGGCA 763
|||||
Db 661 CCAGGATTTGACACAAAGGATGTGAGGAGGCTGTAGGAGCGGCTTGTACCGGGTGGCA 720
QY 764 GTGCCCTCACTGGGCGCAGCCACTACAGAGAGAGTGAAGTCCGAGAGAGAGTGCCTCA 823
|||||
Db 721 GTAGCCTCACTGGGCGCAGCCACTACAGAGAGAGTGAAGTCCGAGAGAGAGTGCCTCA 780
QY 824 ATGAGTGG 831
|||||
Db 781 ATGAGTGG 788

RESULT 5
B0014446/c 772 bp mRNA linear EST 26-MAR-2002
LOCUS UI-H-ED1-ext-9-13-0-UI.s1 NCI_CGAP_ED1 Homo sapiens cDNA clone
DEFINITION IMAGE:5833380 3', mRNA sequence.
ACCESSION B0014446
VERSION B0014446.1 GI:19739347
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 772)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Jose Mercenude
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Seq primer: M13 FORWARD
POLYA=yes.

Db 125 CTCCTGACACTGAGGATGCAACAACCTCTTGAGAGAACCCCTCAGCTGTGGC 184
 QY 2055 GCACATATCTCTGAGCTGCGCCCATACCGCTACTACAGCTCAGCTCCACCCCTG 2114
 Db 185 GCACATATCTCTGAGCTGCGCCCATACCGCTACTACAGCTCAGCTCCACCCCTG 244
 QY 2115 TCACATGAGGCTGACCTGACCTCCACCCCTGTCATACAGCTCAGCTCCACCTG 2174
 Db 245 TCACATGAGGCTGACCTGACCTCCACCCCTGTCATACAGCTCAGCTCCACCTG 304
 QY 2175 CCAGGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2234
 Db 305 CCAGGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 364
 QY 2235 ACTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2294
 Db 365 ACTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 424
 QY 2295 ACTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2354
 Db 425 ACTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 484
 QY 2355 CCCCCCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2414
 Db 485 CCCCCCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 544
 QY 2415 CAGGCAACGAGTGGCCCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2474
 Db 545 CAGGCAACGAGTGGCCCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 604
 QY 2475 CAGGAGGCTGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2534
 Db 605 CAGGAGGCTGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 663
 QY 2535 CCGACACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2594
 Db 664 CCGACA-CTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 721
 QY 2595 TTTGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2654
 Db 722 TTTGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 781
 QY 2655 ACTGGAATTAACATTTTCATTAACATTAACATTAACATTAACATTAACATTAAC 2697
 Db 782 ACTGGAATTAACATTTTCATTAACATTAACATTAACATTAACATTAACATTAAC 824

RESULT 7
 BE563259 846 bp mRNA linear EST 15-AUG-2000
 LOCUS 601335565F1 NIH_MGC_39 Homo sapiens cDNA clone IMAGE:3689369 5',
 DEFINITION mRNA sequence.
 ACCESSION BE563259
 VERSION BE563259.1 GI:9807071
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 846)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabp@remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: LNCM384 Row: a Column: 18

High quality sequence stop: 709.
 FEATURES
 Location/Qualifiers
 1..846
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_1db="NIH_MGC_39"
 /tissue_type="adenocarcinoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
 Site_2: EcoRI; cDNA made by oligo-dT priming.
 directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCAGAG(G). Library constructed
 by Ling Hong in the laboratory of Gerald M. Rubin
 (University of California, Berkeley) using ZAP-cDNA
 synthesis kit (Stratagene) and Superscript II RT (Life
 Technologies)"
 BASE COUNT 172 a 252 c 266 g 156 t
 ORIGIN
 Query Match 25.8%; Score 698.6; DB 10; Length 846;
 Best Local Similarity 96.9%; Pred. No. 1.3e-95;
 Matches 776; Conservative 0; Mismatches 19; Indels 6; Gaps 6;
 QY 474 CCTGCTGCTAGTTCTTACACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 533
 Db 1 CCTGCTGCTAGTTCTTACACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
 QY 534 GGGCGTGATTCCTCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 593
 Db 61 GGGCGTGATTCCTCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
 QY 594 CTGAGTGCACACCGAGTGTACTTATGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 653
 Db 121 CTGAGTGCACACCGAGTGTACTTATGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
 QY 654 GCAAGCCGGATCTTCAAGCCATCTCATCCAGACATGTGTGGCCACACTCCAGTAT 713
 Db 181 GCAAGCCGGATCTTCAAGCCATCTCATCCAGACATGTGTGGCCACACTCCAGTAT 240
 QY 714 GCACCAAGATGTGAGGAGCTCTAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 773
 Db 241 GCACCAAGATGTGAGGAGCTCTAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
 QY 774 CTGGGCGACCCATACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 833
 Db 301 CTGGGCGACCCATACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
 QY 834 GGTATGCGCCAGCTGAGATCTCTGCGGCTCCACAGCCCGAGCTGGCGGGTCTCTAGA 893
 Db 361 GGTATGCGCCAGCTGAGATCTCTGCGGCTCCACAGCCCGAGCTGGCGGGTCTCTAGA 420
 QY 894 ACAGAGAGAGATGAG 953
 Db 421 ACAGAGAGAGATGAG 480
 QY 954 CCTGAGAGATGTCACTTCCAAAGAGATCCGCGAGCTGTGAGCTGCGCTGGGGCTCC 1013
 Db 481 CCTGAGAGATGTCACTTCCAAATAGATCCG-CAGGCTGTGAGCTGCGCTGGGGCTCC 539
 QY 1014 CCTCAGAGATGACGCTGACTTATGACACACAGATGCTGCTGCTGCTGCTGCTGCTG 1073
 Db 540 CCTCAGAGATGACGCTGACTTATGACACACAGATGCTGCTGCTGCTGCTGCTGCTG 598
 QY 1074 CCGAGCTCCCGCATCTCCCGCATCTCCCGCATCTCCCGCATCTCCCGCATCTCCCGCAT 1133
 Db 599 CCGAGCTCCCGCATCTCCCGCATCTCCCGCATCTCCCGCATCTCCCGCATCTCCCGCAT 657
 QY 1134 GGAGAGCTGACAG 1193
 Db 658 GGAGAGCTGACAG 715
 QY 1194 CTTTACCTTGAGCGCTTACCTACCAATGTGCGCTCTGGGATGAGAGAGAGAGAGAG 1253

Db 716 CTTTAACTGAGCGCTTCA-CTACCAATGTGGCTTCTGGATGAGAGTGGCGCA 774
Qy 1254 GCTGCTGCCGACCTGAGAGA 1274
Db 775 CTGTGCGCTGTAGGAGCGCA 795

RESULT 8
BG340220
LOCUS BG340220 932 bp. mRNA linear EST 27-FEB-2001
DEFINITION 602438408F1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4556401 5',
mRNA sequence.
ACCESSION BG340220
VERSION BG340220.1 GI:13146647
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 932)
NIH-MGC http://mgc.nci.nih.gov/
NATIONAL INSTITUTES OF HEALTH, Mammalian Gene Collection (MGC)
UNPUBLISHED (1999)
CONTACT: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L10M1258 row: P column: 02
High quality sequence stop: 630.
Location/Qualifiers
1. 932
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4556401"
/clone_lib="NIH_MGC_46"
/tissue_type="leiomyosarcoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: uterus; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; CDNA made by oligo-dT priming, directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the Laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

BASE COUNT 160 a 284 c 326 g 162 t
ORIGIN

Query Match 25.7%; Score 695.6; DB 12; Length 932;
Best Local Similarity 93.0%; Pred. No. 3.5e-95;
Matches 794; Conservative 0; Mismatches 54; Indels 6; Gaps 6;

Qy 52 GGGCTGCGCCGGTGCAGCCAGTGTCTGCGGCTGCTGCATGAGCCCTGTGTCACAGT 111
Db 1 GGGCTGCGCCGGTGCAGCCAGTGTCTGCGGCTGCTGCATGAGCCCTGTGTCACAGT 60
Qy 112 AGCGCTTGGCCCCGGGAGAGGCGGCTTCACAGCCCGTGGGGCCCTGGAGACAGCGGCT 171
Db 61 AGCGCTTGGCCCCGGGAGAGGCGGCTTCACAGCCCGTGGGGCCCTGGAGACAGCGGCT 120
Qy 172 CAGCAAGAGATGCACTCAGCAAGAGAGAGCTTGGCGGTCTCCGTGGGCTGTCTCTG 231
Db 121 CAGCAAGAGATGCACTCAGCAAGAGAGAGCTTGGCGGTCTCCGTGGGCTGTCTCTG 180
Qy 232 GAGCTGCAAGATGAGAGGAGACATGATGAGAGAGAGCCAGTTCTGAGCCAAACAGAG 291

Db 181 GAGCTGCAAGATGAGAGGAGACATGATGATGAGAGAGGCCAGTTCTGAGCCAAACAGAG 240
Qy 292 AAGGCCCCGAGTGAAGAGAGGCTCCAGCGGGACACAGAGACTTGGGCAAGATCCAG 351
Db 241 AAGGCCCCGAGTGAAGAGAGGCTCCAGCGGGACACAGAGACTTGGGCAAGATCCAG 300
Qy 352 AGTCCCAAGAAACAGAGAGAGAGAGAGGAGAGGAGAGGAGAGGAGAGGAGAGGAGAG 411
Db 301 AGTCCCAAGAAACAGAGAGAGAGAGAGGAGAGGAGAGGAGAGGAGAGGAGAGGAG 360
Qy 412 CCGCAGATGATATCCGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 471
Db 361 CCGCAGATGATATCCGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 419
Qy 472 TACCTGCTGATGTTTCTACAG 531
Db 420 TACCTGCTGATGTTTCTACAG 479
Qy 532 CTGGGCTGATTTCCCTGACAG 591
Db 480 CTGGGCTGATTTCCCTGACAG 539
Qy 592 CTCTGAGTGAACACCAAGGTACTTATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 651
Db 540 CTCTGAGTGAACACCAAGGTACTTATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 599
Qy 652 GGGCAAGAGCGGATCTCAAGCCCATCTCCATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 710
Db 600 GGGCAAGAGCGGATCTTCAAG-CCATCTCCATCAAGAGAGAGAGAGAGAGAGAGAGAGAG 658
Qy 711 ATTGCACCAAGCATGTGAG 769
Db 659 ATTGCACCAAGCATGTGAG 718
Qy 770 TCACCTGGGCGAGCCACTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 829
Db 719 TCAGCTGGG-CAGCACTTCCCGGGGAGAGCTGAGCTGCC-ACAGAGAGTCTTCATGAGAT 776
Qy 830 GAGAGGCTATGGCGCGAGCTGAGAGTCTGCGGCTGCCAGCGCGAGAGAGAGAGAGAGAG 889
Db 777 GCGCGGATGTGCGAGAGCTGAGAGTTCGCGGCTCCAGCCAGAGCTGGGGGCTCGACA 836
Qy 890 CAGACAGAGAGAGAG 903
Db 837 CCGGAGATGAGAGAG 850

RESULT 9
BE869222 726 bp mRNA linear EST 20-OCT-2000
LOCUS BE869222
DEFINITION 601445030F1 NIH_MGC_65 Homo sapiens cDNA clone IMAGE:3849342 5',
mRNA sequence.
ACCESSION BE869222
VERSION BE869222.1 GI:10317998
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 726)
NIH-MGC http://mgc.nci.nih.gov/
NATIONAL INSTITUTES OF HEALTH, Mammalian Gene Collection (MGC)
UNPUBLISHED (1999)
CONTACT: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L10M9566 row: K column: 07

Db 419 GATGCTGCTGCTGTGTGGCAGACGGGACCGAGCTCCGCACTCTTCCCCACCTCTACCT 478
QY 1107 GGGCTCAGAGTGGAGCGCAGCAACCTGGAGAGCTGCAGAGAACAGGGTCAACCCACAT 1166
Db 479 GGGCTCAGAGTGGAGCGCAGCAACCTGGAGAGCTGCAGAGAACAGGGTCAACCCACAT 538
QY 1167 CTTGAACATGGCCCGGAGATTCAGCACTTCACTGAGCCCTTCACTTACCTACCAATGT 1226
Db 539 CTTGAACATGGCCCGGAGATTCAGCACTTCACTGAGCCCTTCACTTACCTACCAATGT 598
QY 1227 GCGCCTCTGGAGTGAAGAGTCCGCCAGCTGCTGCCGCACTGGAAGAGACGACCCGCTT 1286
Db 599 GCG-CTCTGGAGTGAAGAGTCCGG-CCAGCTGCTGC--GACTGGAAGAGACGACCCGCTT 654
QY 1287 CATTGA-GGCTGCAAGAGACAGGCGACCCAGCTGCTGCTGCACATGCAAGATGGGCTCA 1345
Db 655 CATTGAAGGCTTCAAGAGACAGGCGACCCAGCTGCTGCTGCACATGCAAGATGGGCTCA 714
QY 1346 GCGCCTCAGCGCCACAGTGTGCTGCTATGCTATGCAAGCAGTACGATGACCTGAGAC 1405
Db 715 GCGGCTCAGCGG-CACAGGGCTGGCTATGCTATGCAAGCAGTACGATGACCTGAGAC 773
QY 1406 AGCCCTCGGC 1416
Db 774 GCGCTGGGAC 784

RESULT 11

Bg753072

LOCUS 60273241F1 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:4875562 5',
DEFINITION mRNA sequence.

ACCESSION Bg753072

VERSION Bg753072.1 GI:14063725

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

1 (bases 1 to 663)

NIH-MGC http://mgi.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: L1CM1757 row: b column: 11

High quality sequence stop: 663.

FEATURES

source
location/Qualifiers

1..663

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone IMAGE:4875562"

/clone_lib="NIH_MGC_43"

/tissue_type="normal pigmented retinal epithelium"

/lab_host="DH10B (phage-resistant)"

/note="Organ: eye; Vector: pORF7; Site_1: XhoI; Site_2:

EcoRI; cDNA made by oligo-dT priming. Directionally

cloned into EcoRI/XhoI sites using the following 5'

adaptor: GGCACGAG(G). Library constructed by Ling Hong

in the laboratory of Gerald M. Rubin (University of

California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies).

Note: this is a NIH_MGC library. !"

Query Match 24.5%; Score 662; DB 12; Length 663;
Best Local Similarity 100.0%; Pred. No. 4,66-90;
Matches 662; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 AGACTGTCCGCGGGGTTGAGGAAAGGGCGTCCCGGTGCAGCCAGTGTCTCGG 84
Db 2 AGACTGTCCGCGGGGTTGAGGAAAGGGCGTCCCGGTGCAGCCAGTGTCTCGG 61
QY 85 CCTGGCTCCATGGCCCTGTGTACAGTACAGTCCCTCCCGCGGCGAGCGCCCTCCAG 144
Db 62 CCTGGCTCCATGGCCCTGTGTACAGTACAGTCCCTCCCGCGGCGAGCGCTCCAG 121
QY 145 CCCGTGGGCGCTTGGAGCAGCAGCGGTCCAGCAAGAGTGCAGTCCAGCAAGAGCAGAG 204
Db 122 CCCGTGGGCGCTTGGAGCAGCAGCGGTCCAGCAAGAGTGCAGTCCAGCAAGAGCAGAG 181
QY 205 TTTGCGGTCTCCGTGGGCGTGTCTGTGAGTCCAGATGAGAGGAGCAATGATGATGA 264
Db 182 TTTGCGGTCTCCGTGGGCGTGTCTGTGAGTCCAGATGAGAGGAGCAATGATGATGA 241
QY 265 GCAGAGCGCGCTTGGAGCAGCAGAGAGGCGCCGAGTGGAGAGAGAGTCCAGCGGGGAC 324
Db 242 GCAGAGCGCGCTTGGAGCAGCAGAGAGGCGCCGAGTGGAGAGAGAGTCCAGCGGGGAC 301
QY 325 CAGACAGACTTGGGCGAAGATCCAGAGTCCAGCAAGAGAGAGAGAGAGAGAGAGAG 384
Db 302 CAGACAGACTTGGGCGAAGATCCAGAGTCCAGCAAGAGAGAGAGAGAGAGAGAGAGAG 361
QY 385 CTGACACTTGTGTGACAGTGTGAGGCGCGAGAGATGACATCCGCTGAGCGCCAGCTG 444
Db 362 CTGACACTTGTGTGACAGTGTGAGGCGCGAGAGATGACATCCGCTGAGCGCCAGCTG 421
QY 445 GAGGCGCGCGCGCTCCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 504
Db 422 GAGGCGCGCGCGCTCCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 481
QY 505 GGTCTGAGCAGAGATGAGAGGCTCTCTGAGGCGTGGATTTCTGAGAGAGAGCTCCCG 564
Db 482 GGTCTGAGCAGAGATGAGAGGCTCTCTGAGGCGTGGATTTCTGAGAGAGAGCTCCCG 541
QY 565 AGCTGACCTTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 624
Db 542 AGCTGACCTTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 601
QY 625 GAGGCGCGCTTGGAGCAGTGTGAGGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAG 684
Db 602 GAGGCGCGCTTGGAGCAGTGTGAGGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAG 661
QY 685 CA 686
Db 662 CA 663

RESULT 12

BM910493

LOCUS BM910493 1004 bp mRNA linear EST 12-MAR-2002
DEFINITION AGENCOURT_6703004 NIH_MGC_98 Homo sapiens cDNA clone IMAGE:5453997
5', mRNA sequence.

ACCESSION BM910493

VERSION BM910493.1 GI:19360872

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

1 (bases 1 to 1004)

NIH-MGC http://mgi.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

BASE COUNT
ORIGIN

124 a 205 c 231 g 103 t

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 11, 2003, 00:00:35 ; Search time 156.754 Seconds
(without alignments)
5290.174 Million cell updates/sec

Title: US-09-761-640-1

Perfect score: 2704

Sequence: 1 cgtctctctctgtctctgcg.....aaaaaaaaaaaaaaaaaaaaa 2704

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued_Patents_NA.*
1: /cgn2.6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2.6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2.6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2.6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2.6/ptodata/2/ina/PTCUTS_COMB.seq:*
6: /cgn2.6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	104.4	3.9	1390	US-09-704-139-1	Sequence 1, Appl1
2	103.4	3.8	531	US-09-704-139-3	Sequence 3, Appl1
3	88	3.3	1238	US-08-530-290-11	Sequence 11, Appl1
4	78.4	2.9	1987	US-08-990-379-1	Sequence 1, Appl1
5	78.4	2.9	1993	US-08-990-379-2	Sequence 2, Appl1
6	61.6	2.3	1729	US-09-045-973-6	Sequence 6, Appl1
7	60	2.2	7218	US-08-232-463-14	Sequence 14, Appl1
8	56.2	2.1	594	US-09-163-833-3	Sequence 3, Appl1
9	55.4	2.0	912	US-09-163-833-1	Sequence 1, Appl1
10	52.6	1.9	1691	US-09-013-881-12	Sequence 12, Appl1
11	52.6	1.9	216	US-08-530-290-9	Sequence 9, Appl1
12	50.2	1.9	861	US-07-988-273-1	Sequence 1, Appl1
13	50.2	1.9	861	PCR-US93-12019-1	Sequence 1, Appl1
14	48.6	1.8	477	US-09-135-994-1	Sequence 1, Appl1
15	48.2	1.8	237	US-08-530-290-8	Sequence 8, Appl1
16	46.8	1.7	1926	US-09-249-585A-2	Sequence 2, Appl1
17	46.8	1.7	2580	US-09-050-863-2	Sequence 2, Appl1
18	46.8	1.7	2580	US-09-358-081-2	Sequence 2, Appl1
19	46.8	1.7	5452	US-09-130-114-1	Sequence 1, Appl1
20	46.8	1.7	5600	US-08-910-647-1	Sequence 1, Appl1
21	46.8	1.7	9600	US-09-620-925-1	Sequence 1, Appl1
22	46.8	1.7	10596	US-07-884-811-15	Sequence 15, Appl1
23	46.8	1.7	10596	US-07-885-971-15	Sequence 15, Appl1
24	46.8	1.7	10596	US-08-087-783A-15	Sequence 15, Appl1
25	46.8	1.7	10596	US-08-194-088B-15	Sequence 15, Appl1
26	46.8	1.7	10596	US-08-194-087-15	Sequence 15, Appl1
27	46.8	1.7	10596	PCR-US93-04648-15	Sequence 15, Appl1

C 28	46.2	1.7	1926	US-09-249-585A-2	Sequence 2, Appl1
C 29	46.2	1.7	2580	US-09-050-863-2	Sequence 2, Appl1
C 30	46.2	1.7	2580	US-09-359-081-2	Sequence 2, Appl1
C 31	46.2	1.7	5452	US-09-130-114-1	Sequence 1, Appl1
C 32	46.2	1.7	9600	US-08-910-647-1	Sequence 1, Appl1
C 33	46.2	1.7	9600	US-09-620-925-1	Sequence 1, Appl1
C 34	46.2	1.7	10596	US-07-884-811-15	Sequence 15, Appl1
C 35	46.2	1.7	10596	US-07-885-971-15	Sequence 15, Appl1
C 36	46.2	1.7	10596	US-08-087-783A-15	Sequence 15, Appl1
C 37	46.2	1.7	10596	US-08-194-088B-15	Sequence 15, Appl1
C 38	46.2	1.7	10596	US-08-194-087-15	Sequence 15, Appl1
C 39	46.2	1.7	10596	PCR-US93-04648-15	Sequence 15, Appl1
C 40	45.6	1.7	2214	US-08-864-038A-1	Sequence 1, Appl1
C 41	45.6	1.7	3331	US-08-864-038A-2	Sequence 2, Appl1
C 42	45.6	1.7	3331	US-08-864-038A-4	Sequence 4, Appl1
C 43	45.4	1.7	1358	US-09-013-881-11	Sequence 11, Appl1
C 44	44	1.6	7218	US-08-232-463-14	Sequence 14, Appl1
C 45	43.8	1.6	35828	US-09-449-218D-17	Sequence 17, Appl1

ALIGNMENTS

```

RESULT 1
US-09-704-139-1
; Sequence 1, Application US/09704139
; Patent No. 6420153
; GENERAL INFORMATION:
; APPLICANT: Kapeller, Rosanna
; TITLE OF INVENTION: 18232, A NOVEL DUAL SPECIFICITY PHOSPHATASE AND USES THEREOF
; FILE REFERENCE: 10448-018001
; CURRENT APPLICATION NUMBER: US/09/704,139
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: US 60/185,772
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 1390
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (329)..(859)
; NAME/KEY: misc.feature
; LOCATION: (1)..(1390)
; OTHER INFORMATION: n = a, t, g, or c
US-09-704-139-1

```

Query Match 3.9%; Score 104.4; DB 4; Length 1390;
Best Local Similarity 51.3%; Pred No. 7.2e-15;
Matches 243; Conservative 0; Mismatches 231; Indels 0; Gaps 0;

```

QY 1084 CGCATCTTCCCGCCACCTTACCTGAGTGCAGTGCAGCAAGCAACCTGAGAGAGCTG 1143
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 389 CGCGTGCAGCCCTCACTCTTCTTCCGAGAGCGCGGCGGCGGAGGAGGAGCGAGCTG 448
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1144 CAGAGCAAGAGGCTACCCACATCTTGAACATGCGCCCGGAGATGACACTTCTACCT 1203
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 449 GCGCGGCGGAGGATCAGCTGTGCGTCAACGTCCTCCGACACAGCCCGCGCGCGG 508
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1204 GAGCGCTTACCTACACATGAGCGGCTCTGAGATGAGAGAGCGGCGGCGGAGCTG 1263
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 509 CCGCGGCTGCGAGAGCTGCGGCTGCCGCTGTCGACAGCCCGCTGAGACCTGCTGCGG 568
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1264 CACTGGAAGAGAGAGCGACCGCTTATGAGCTGCAAGAGCAAGGCGACCGACGCTG 1323
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 569 CACGTGAGACCGACGCGCGCCGACGAGGAGCGGCGGCGGCGGCGGCGGCGGCTG 628
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1324 GTCCACTGCAAGATGGCGCTGACGCGGCTAGCGGCGACAGTCTGCTATGCGCATGAAG 1383
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 629 GTCTACTGCAAGAGAGCGCGGCGGCTGCGGCGGCTGCGACGCGCTACTCATCTGCGG 688
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

QY 1384 CAGTAGAATGAGCGCTGAGCAGGCGCTGCGCCAGCTGACAGAGCTCCGCCATCCG 1443
DB 689 CACCGCGGCTCAGCTGAGCGGCTTCCAGATGTTGAAGAGCGCTCCCGGTAGCA 748
QY 1444 CGCCCAACCCCTGCTCTGCGGCGAGCTGCTACCAAGGCGATCTGTAGCGGCGACA 1503
DB 749 GAACCGAAGCCGGGCTTCTGCTGCTCAGCTCCAGAACTATGAGAGAGGCGCTCCAGGCGCGAG 808
QY 1504 ACCTAGAGGTGGTGGGAGAGAGAGTGTGATGATGGAAGAGAGCGCAGCAGC 1557
DB 809 TCCTGCTGCGAGGAGAGGCGCCAGCTTAGGCTTGAGGCTTGAGGCTTGAAAGC 862

RESULT 2
US-09-704-139-3
; Sequence 3, Application US/09704139
; Patent No. 6420153
; GENERAL INFORMATION:
; APPLICANT: Kapeller, Rosanna
; TITLE OF INVENTION: 18232, A NOVEL DUAL SPECIFICITY PHOSPHATASE AND USES THEREFOR
; FILE REFERENCE: 10448-018001
; CURRENT APPLICATION NUMBER: US/09/704,139
; CURRENT FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: US 60/185,772
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 3
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-704-139-3

Query Match 3.8%; Score 103.4; DB 4; Length 531;
Best Local Similarity 52.7%; Pred. No. 8,1e-15;
Matches 224; Conservative 0; Mismatches 201; Indels 0; Gaps 0;

QY 1084 CGCATCTTCCCGCCACCTCTACCTGAGCTGAGTGAAGCGCAGCAAACTGAGAGAGCTG 1143
DB 61 CGCGTCGCGCGCTCAGCTCTCTCGGGAGCGCGCAGCGCGGCGCGAGAGAGAGCTG 120
QY 1144 CAGAGAACAGGCTACCCACATCTTGAACATGGCCCGGAGATGTGACAACTTCTACCT 1203
DB 121 GCGCGCGCGGAGTCAAGCTGTGCTCAAGCTCTCCCGCAGCAGCGCGCGCGCGAGAG 180
QY 1204 GAGCGCTTCACTTACCAATGTGCGCTCTGGGATGAGAGAGTGGCGCCAGCTGCGCG 1263
DB 181 CCGCGCTGCGAGAGCTGCGCGCTGCGCTGCTGAGAGAGAGCGCGCTGAGAGAGCTG 240
QY 1264 CACTGGAAGAGAGCAGCGCTTCTTATGAGGCTGCAAGAGCAGAGGCGACCCAGTGTG 1323
DB 241 CACCTGAGCGCCAGCTGCGCGCGCAGTGAAGAGCGCGCGCGCGCGCGCGCTGCTA 300
QY 1324 GTTCATGCAAGATGGGCGTCAAGCGCTGCGCGCGCAGCAATGCTGTGCTATGCAATGA 1383
DB 301 GTTCTACTGAAAGAGCGCGCGCGCTGCGCGCTGCGCGCTGCAAGCGCTGCTGCTGCGG 360
QY 1384 CAGTAGAATGAGCGCTGAGCAGGCGCTGCGCGCAGCTGAGAGAGCTCCGCCATCGCC 1443
DB 361 CACCGCGGCTCAGCTGCGGAGAGCGCTTCCAGATGTTGAAGAGCGCTCCCGGTAGCA 420
QY 1444 CGCCCAACCCCTGCTCTGCGGCGAGCTGAGATCTTACCAAGGCGATCTGTAGCGGCGACA 1503
DB 421 GAACCGAAGCCGGGCTTCTGCTGCTCAGAGCTCCAGAACTATGAGAGAGGCGCTCCAGGCGCAG 480
QY 1504 ACCTG 1508
DB 481 TCCTG 485

RESULT 3
US-08-530-290-11
; Sequence 11, Application US/08530290

; Patent No. 5958721
; GENERAL INFORMATION:
; APPLICANT: Marshall, Christopher John
; APPLICANT: Ashworth, Alan
; APPLICANT: Hughes, David Anthony
; TITLE OF INVENTION: Methods for Screening of Substances for
; TITLE OF INVENTION: Therapeutic Activity and Yeast for Use Therein
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/530,290
; FILING DATE: 14-DEC-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/GB94/00694
; FILING DATE: 31-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9402573.1
; FILING DATE: 10-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9307250.2
; FILING DATE: 07-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 084611-000000US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1238 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-530-290-11

Query Match 3.3%; Score 88; DB 2; Length 1238;
Best Local Similarity 50.5%; Pred. No. 4,1e-11;
Matches 214; Conservative 0; Mismatches 210; Indels 0; Gaps 0;

QY 1087 ATCTTCCCGCCACCTCTACCTGAGGCTCAGAGTGAAGCAGCAAACTGAGAGAGCTGAG 1146
DB 648 ATCTTCCCTCTCTTACTTACCTGCGAGTGTCTACCATGCTCCCGGAGAGAGAGCTGGAG 707
QY 1147 AGAAGAGGCTACCCACATCTTGAACATGCGCGGAGATGTGACAACTTCTTACCTGAG 1206
DB 708 GCGCTGAGCATACAGGCTGTGTAATGTCTCTCGAGACTGCCCAACCACTTTGAAGA 767
QY 1207 CGCTTCACTTACCAATGTGCGCTCTGGGATGAGAGAGTGGCGCCAGCTGCTGCGCGAC 1266
DB 768 CACTATCTAGTACAAGTGCATCCAGTGAAGATTAACCAACAAGGCGCAGATCAGCTCTG 827
QY 1267 TGAAGAGAGAGCAGCGCTTCTTATGAGGCTGCAAGAGCAGACAGCGCAGCTGTGTC 1326
DB 828 TTCAATGGAAGCATAGATGATGATGCTGTAAGAGACTGCGCTGCGCGCTGCTGCTG 887
QY 1327 CACTGCAAGATGGCGCTCAGCGCTGAGCGCGCAGCAAGTGTGCTATGCAATGAAGAGAG 1386
DB 888 CACTGCGAGCGGCGCATCTGCGGCTGCGCGCAGCAGCTGCTGCTGCTTACCTGATGATGAAG 947

OY 1387 TAGCAATGACGCTTGAGACAGGCCCTGCGCCACGTGACAGAGATCCGGCCCATCGCCGC 1446
 DB 948 AACGGGTGAGGCTGAGAGAGCCCTTCAGATTCGTTAAGACAGCCGACGATCATCTCG 1007
 OY 1447 CCCAACCTGCTTCTGCTGCGCCAGCTGCAGATCTACACGGGATCTGACGGCCAGAAC 1506
 DB 1008 CCCAACTTCAGCTTCATGCGGACGCTGCTGAGTTCAGTCCAGAGTCTGCGCACGTC 1067
 OY 1507 TGAG 1510
 DB 1068 TGTG 1071

RESULT 4

US-08-990-379-1
 : Sequence 1, Application US/08990379
 : Patent No. 5998188
 : GENERAL INFORMATION:
 : APPLICANT: Stork, Philip J
 : APPLICANT: Miska-Press, Anita
 : TITLE OF INVENTION: Mitogen Activated Protein Kinase Phosphatase cDNAs and
 : TITLE OF INVENTION: Their Biologically Active Expression Products
 : FILE REFERENCE: 4104-000322USA
 : CURRENT APPLICATION NUMBER: US/08/990, 379
 : CURRENT FILING DATE: 1997-12-15
 : EARLIER APPLICATION NUMBER: PCT/US96/10402
 : EARLIER FILING DATE: 1996-06-14
 : EARLIER APPLICATION NUMBER: 60/000, 263
 : EARLIER FILING DATE: 1995-06-16
 : NUMBER OF SEQ ID NOS: 19
 : SOFTWARE: Patentln Ver. 2.0
 : SEQ ID NO 1
 : LENGTH: 1987
 : TYPE: DNA
 : ORGANISM: Rattus norvegicus
 : US-08-990-379-1

Query Match 2.9%; Score 78.4; DB 2; Length 1987;
 Best Local Similarity 49.1%; Pred. No. 8e-09;
 Matches 208; Conservative 0; Mismatches 216; Indels 0; Gaps 0;

OY 1087 ATCTTCCCCACCTTCTACCTGGGCTCAGAGTGAAGCAGCAACCTGAGAGAGCTGCAG 1146
 DB 969 ATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1028
 OY 1147 AGAAGACGGGTACCCACATCTTGAACATGGCCCGGAGATGAGAACTTCAACCTGAG 1206
 DB 1029 GCCCTGGGATACGCGCTCTACTGATGTCTCTCAGACTGCCCAATCATTGAGGGA 1088
 OY 1207 CGCTTACCTACACATATGTGGCTTGGAGTGAAGATCGGCGCAGCTGCTGCGCAC 1266
 DB 1089 CATTCACAGTACAAATGATGATCCCGGTAGAGATTAACCAAGGCTGACATGCTCTGG 1148
 OY 1267 TGAAGAGAGACGACCGCTTCAATGAGCTGCAAGACACAGGAGCCACCTGCTGCTG 1326
 DB 1149 TTCATGAGAGCCATGAAATATAGACAGCATGAGGAGCTGCGGAGGCGAGTGTGTT 1208
 OY 1327 CACTGCAAGATGGGCTGAGCGGCTGAGGCGCACAGTGTGCTATGCTGTAAGAG 1386
 DB 1209 CACTGCGAGGCGGATCTTATGATCAGCAGCAGATCTGCTGCTTACCTGATGATGAG 1268
 OY 1387 TAGCAATGACGCTTGAGACAGGCCCTGCGCCACGTGACAGAGCTTCGGCCCATCGCCGC 1446
 DB 1269 AAAAGGAGAGGCTGAGAGAGGCTTTCAGTTCGTCAGACAGCGCGAGAGCATCATCTCG 1328
 OY 1447 CCCAACCTGCTTCTGCTGCGCCAGCTGCAGATCTACACGGGATCTGACGGCCAGAAC 1506
 DB 1329 CCCAACTTCAGCTTCATGCGGACGCTGCTGAGTTCAGTCTCAGTGTCTCACACAGTCC 1388
 OY 1507 TGAG 1510
 DB 1389 TGTG 1392

RESULT 5
 US-08-990-379-2
 : Sequence 2, Application US/08990379
 : Patent No. 5998188
 : GENERAL INFORMATION:
 : APPLICANT: Stork, Philip J
 : APPLICANT: Miska-Press, Anita
 : TITLE OF INVENTION: Mitogen Activated Protein Kinase Phosphatase cDNAs and
 : TITLE OF INVENTION: Their Biologically Active Expression Products
 : FILE REFERENCE: 4104-000322USA
 : CURRENT APPLICATION NUMBER: US/08/990, 379
 : CURRENT FILING DATE: 1997-12-15
 : EARLIER APPLICATION NUMBER: PCT/US96/10402
 : EARLIER FILING DATE: 1996-06-14
 : EARLIER APPLICATION NUMBER: 60/000, 263
 : EARLIER FILING DATE: 1995-06-16
 : NUMBER OF SEQ ID NOS: 19
 : SOFTWARE: Patentln Ver. 2.0
 : SEQ ID NO 2
 : LENGTH: 1993
 : TYPE: DNA
 : ORGANISM: Rattus norvegicus
 : US-08-990-379-2

Query Match 2.9%; Score 78.4; DB 2; Length 1993;
 Best Local Similarity 49.1%; Pred. No. 8.1e-09;
 Matches 208; Conservative 0; Mismatches 216; Indels 0; Gaps 0;

OY 1087 ATCTTCCCCACCTTCTACCTGGGCTCAGAGTGAAGCAGCAACCTGAGAGAGCTGCAG 1146
 DB 975 ATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1034
 OY 1147 AGAAGACGGGTACCCACATCTTGAACATGGCCCGGAGATGAGAACTTCAACCTGAG 1206
 DB 1035 GCCCTGGGATACGCGCTCTACTGATGTCTCTTCAAGTGTGCGCCCAATCATTGAGGGA 1094
 OY 1207 CGCTTACCTACACATATGTGGCTTGGAGTGAAGATCGGCGCAGCTGCTGCGCAC 1266
 DB 1095 CATTCACAGTACAAATGATGATCCCGGTAGAGATTAACCAAGGCTGACATGCTCTGG 1154
 OY 1267 TGAAGAGAGACGACCGCTTCAATGAGGCTGCAAGACACAGGAGCCACCTGCTGCTG 1326
 DB 1155 TTCATGAGAGCCATGAAATATAGACAGCATGAGGAGTGTGCGGAGAGTGTGTT 1214
 OY 1327 CACTGCAAGATGGGCTGAGCGCTCAGCGGCGCACAGTGTGCTTATGCTGTAAGAG 1386
 DB 1215 CACTGCGAGGCGGATCTTATGATCAGCAGCAGATCTGCTGCTTACCTGATGATGAG 1274
 OY 1387 TAGCAATGACGCTTGAGACAGGCCCTGCGCCACGTGACAGAGCTTCGGCCCATCGCCGC 1446
 DB 1275 AAAAGGAGAGGCTGAGAGAGGCTTTCAGTTCGTCAGACAGCGCGAGAGCATCATCTCG 1334
 OY 1447 CCCAACCTGCTTCTGCTGCGCCAGCTGCAGATCTACACGGGATCTGACGGCCAGAAC 1506
 DB 1335 CCCAACTTCAGCTTCATGCGGACGCTGCTGAGTTCAGTGTCTCAGTGTCTCACACAGTCC 1394
 OY 1507 TGAG 1510
 DB 1395 TGTG 1398

RESULT 6
 US-09-045-973-6
 : Sequence 6, Application US/09045973
 : Patent No. 6165767
 : GENERAL INFORMATION:
 : APPLICANT: Lal, Preeti
 : APPLICANT: Yue, Henry
 : APPLICANT: Corley, Neill C.
 : APPLICANT: Guejter, Karl J.
 : APPLICANT: Baughn, Mariah
 : TITLE OF INVENTION: PROTEIN PHOSPHATASE RELATED MOLECULES

```

NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/045,973
FILING DATE: Filed Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0491 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1729 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRSTW0116
CLONE: 3041794

```

	Query Match	Similarity	2.3%	Score	61.6:	DB	Length	1729:
	Best Local	Similarity	46.8%:	Pred.	No. 5.6e-05:			
	Matches	193:	Conservative	0:	Mismatches	219:	Indels	0:
							Gaps	0
QY	1082	CCCCGACTCTTCCGCCACACCCTACTCGGGCTCAGAGTGGAACGCACGCAAACTGGAGAGAC	1141					
Db	574	CTCAATATCACCTCTCTCTATTCTTCGTGGGAGAGGACAGTGTTGGGCTTCACATCGGACCTCC	633					
OY	1142	TGCAGAAGNACAGGGTGCACCCCATCTTATAACATGGCCCGGGAGATGTACAACCTTTACC	1201					
Db	634	TCCAGGCTCGTGGCATCACCTGCATTTGTTAAATGCTAACATTTGATGATCCCTAATTTTAACCT	693					
OY	1202	CTGAGCGGCTTCACTACACATGAGCGGCTCTGGGATGAGGAGTGGGCGACCTGCTGC	1261					
Db	694	GGCCCCAATTTTAGTATGTATTAAAGTCCCTCTGTGGCTGACATGCCCGCATGCCCATTTGGAC	753					
OY	1262	CGCACGTGNAAGGAGACGACCCGCTTCATTAGAGGCTGCAAGACACAGGGGACCCACCTGC	1321					
Db	754	TCTTACTTTGACACCGTGGCTGACACAGATCCACAGTGTGAGACAGAGGACACGGGGCCACCT	813					
OY	1322	TGGTCACTGCAAGATGGGCGCTCAGCCGCTCAGCGGCCACAGTGTCTGGCCTATGCATGA	1381					
Db	814	TGGTCACTGCTCTCTGAGGGGTGAGCGGCTCAGCCACGCTGTGTATCGGGTACTCCTGTA	873					
OY	1382	AACAGTAGCAATGACGCGCTGGAGACGAGCCCTGGCCACAGGACGAGAGCTCCGGCCCATCT	1441					
Db	874	AATTTCCACACAGCTGTGCTGCTGGAGGCTTACAACTGGGGTGAAGCCCGGCGACCTGTCA	933					
OY	1442	CCGCCCCAACCCCTGGCTTCTGTGGCCACCTGCAGATCTACCAAGGGGCATCTT	1493					
Db	934	TGAGGCGCCACAGCTAGGCTTCTGTGAGGCACTGATPACACTACGAGGCCACGCT	985					

```

US-08-232-463-14
: Sequence 14, Application US/08232463
: Patent No. 5670367
:
: GENERAL INFORMATION:
: APPLICANT: DORNER, F.
: APPLICANT: SCHEFFLINGER, F.
: APPLICANT: FALKNER, F. G.
: TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
: NUMBER OF SEQUENCES: 52
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Foley & Lardner
: STREET: 1800 Diagonal Road, Suite 500
: CITY: Alexandria
: STATE: VA
: COUNTRY: USA
: ZIP: 22313-0299
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentln Release #1.0, Version #1.25
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/232,463
:
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/07/935,313
: FILING DATE:
: APPLICATION NUMBER: EP 91 114 300.6
: FILING DATE: 26-AUG-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: BENT, Stephen A.
: REGISTRATION NUMBER: 29,768
: REFERENCE/DOCKET NUMBER: 30472/114 IMMU
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703)836-9300
: TELEFAX: (703)683-4109
:
: TELEX: 899149
:
: INFORMATION FOR SEQ ID NO: 14:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 7218 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
:
: IMMEDIATE SOURCE:
: CLONE: PT29pt-Fls
:
US-08-232-463-14

```

[illegible]

QY	1389	CGATGACAGCCTGGAGCAGGCCCCCTGGGCCACATGACAGAGCTGCGGCCCATGCCCCGCC	1448
Db	912	CCACCTTACCTCGTGGAGGCCATCAAGAAAGTCAAGAGCCAGGAGTAT--CCG	968
QY	1449	CAACCCGTGCTCCTGGGCCAGCTGAGATCTTACCGGGGCATCTGCAGCGCCAGAACCTG	1508
Db	969	CAACCGGGGCTCTCTGAGGAGCTCTCTGGGCCCTGGACCCGACAGGCTGCGGCAAGGCTTGGA	1028
QY	1509	AGGCTGCTGGGAGAGAGAGCTGT	1533
Db	1029	AGCATGAGGGAGGGGAGAGAGCT	1053

```

US-08-530-290-9
; Sequence 9, Application US/08530290
; Patent No. 5958721
; GENERAL INFORMATION:
; APPLICANT: Marshall, Christopher John
; APPLICANT: Ashworth, Alan
; APPLICANT: Hughes, David Anthony
; TITLE OF INVENTION: Methods for Screening of Substances for
; TITLE OF INVENTION: Therapeutic Activity and Yeast for Use Therein
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/530,290
; FILING DATE: 14-DEC-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/GB94/00694
; FILING DATE: 31-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9402573.1
; FILING DATE: 10-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9307250.2
; FILING DATE: 07-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 084611-00000005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEO ID NO.: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 216 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; US-08-530-290-9

Query Match 1.9%; Score 52.6; DB 2; Length 216;
Best Local Similarity 53.7%; Pred. No. 0.0028;
Matches 109; Conservative 0; Mismatches 94; Indels 0; Gaps

QY 1262 CGCACTGGAAGAGACGACCGCTTCATTGAGCGTCGAAAGACACAGCGGACCGACCGTC 1322
14 CCCACTTTCACAAACCAATTGATTTCATTGTACTGCGTCGACGGAAGAGAGACGCAAGGTC 73

```

Oy	1322	TGCTCCACTGCAGATGGCGGTCAAGCCGCTCAGGCCACAGTGGCGGCTATTGCCATGA	1381
Dd	74	TATGTCACCTGTGAAGGTGGGGTTCTCAGAGGTCAACCCACCTTGCATATGGGTACTCTCATGA	133
Oy	1382	AGCAGTACGATGAGCGCTCGTAGCACAGGCCCTCGGCCACGTGCAGAGAGCTCCGAGCCATCG	1441
Dd	134	AGACCACAGCAGTTCCGCTCATAAGAGAGGCGCTTGACATCGTCAAGCAGAAGAGAGAACGTGA	193
Oy	1442	CCGCCCCAACCCTGGGCTTCTCTG	1464
Dd	194	TTCTCTCCCAACTTTGGCTTTATG	216

```

      RESULT 12
      US-07-988-273-1
      ; Sequence 1, Application US/07988273
      ; Patent No. 5512434
      ; GENERAL INFORMATION:
      ; APPLICANT: AARONSON, Stuart A.
      ; APPLICANT: BOTTARO, Donald P.
      ; APPLICANT: ISHIBASHI, Toshio
      ; APPLICANT: MIKI, Toru
      ; TITLE OF INVENTION: EXPRESSION CLONING OF A HUMAN
      ; TITLE OF INVENTION: PHOSPHATASE
      ; NUMBER OF SEQUENCES: 7
      ; CORRESPONDENCE ADDRESS:
      ; ADDRESSEE: Foley & Lardner
      ; STREET: 3000 K Street, N.W., Suite 500
      ; CITY: Washington, D.C.
      ; COUNTRY: USA
      ; ZIP: 20007-5109
      ; COMPUTER READABLE FORM:
      ; MEDIUM TYPE: Floppy disk
      ; COMPUTER: IBM PC compatible
      ; OPERATING SYSTEM: PC-DOS/MS-DOS
      ; SOFTWARE: Patent In Release #1.0, Version #1.25
      ; CURRENT APPLICATION DATA:
      ; APPLICATION NUMBER: US/07/988, 273
      ; FILING DATE: 19921214
      ; CLASSIFICATION: 435
      ; ATTORNEY/AGENT INFORMATION:
      ; NAME: BENT, Stephen A.
      ; REGISTRATION NUMBER: 29,768
      ; REFERENCE/DOCKET NUMBER: 40399/182 NIHD
      ; TELECOMMUNICATION INFORMATION:
      ; TELEPHONE: (202)672-5300
      ; TELEFAX: (202)672-5399
      ; TELEX: 904136
      ; INFORMATION FOR SEQ ID NO: 1:
      ; SEQUENCE CHARACTERISTICS:
      ; LENGTH: 861 base pairs
      ; TYPE: NUCLEIC ACID
      ; STRANDEDNESS: double
      ; TOPOLOGY: linear
      ; FEATURE:
      ; NAME/KEY: CDS
      ; LOCATION: 29..586
      ; US-07-988-273-1

Query Match          1.9%; Score 50.2; DB 1; Length 861;
Best Local Similarity 61.6%; Pred. No. 0.018;
Matches 98; Conservative 0; Mismatches 58; Indels 3; Gaps 1

QY 1314 CCACGCTGCAGTGGTCACCTGCAAGATGGCGCCTCAGCCGCTCAGCGGCCACACAGTGCTGCTCTA 1373
    ||| ||||| ||||| ||||| ||| ||| ||||| ||| ||| |||||
Db   382 CCGGGGTACTGCTGCTCACCTGCGCGGAAGATTATAGCCGCTCCCCAACGCTTAGTTATGCGCTTA 441

QY 1374 TGGCATTAAGACACATGACATGACATGAGCGCTGAGACAGGCGCTTGCGCCACATGCGAGAGCTCCG 1433
    ||||| ||| ||| ||| ||| ||| ||||| ||| ||| ||||| ||| |||
Db   442 CCTCATGATGCGGACAAGAAGATGAGACGCTCAAGTCTGCGCTGAGCATGTGATGAGGACGAACGG 501

QY 1434 GCCCATGCGCCCGCCCAACCCTTGCTTCTGCGCCCAAGCT 1472

```


Matches 119; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

```

Oy 1087 ATCTTCCCCCACCCTCTACCTGGGCTCAGAGTGAACGACGCAACCTGGAGAGCTGCAG 1146
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   1  ATCTTGCCCTTACCTGTTCCTGGGCTGCACTCCTCTGACACCTGGAGGGCTGCAG 60
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Oy 1147 AGGACAGGGGTCACCCACATCTTGAACATGGCCCGGAGATTGACACTTCTACCCCTGAG 1206
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   61 GCGCTGGGATACAGAGCGCTCTCAACGCTGCCGACGCTGCCCAACCACTTTGAGGGC 120
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Oy 1207 CGCTTACCTACCAATATGTGGGCTGTGGATGAGAGATCGGCCAGCTGCTGCCGCAC 1266
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   121 CTTTCCGCTACAGAGATATCCCTGTGAGGACCAACAGATGTGAGATCAGTGCCTGG 180
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Oy 1267 TGGAGGAGAGACGACCGCTTCATTTGAGGCTGCAAGAGACAGGACCCACGCTGCTG 1323
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   181 TTCCAGGAGGCTCATAGGCTTCACTGTGGTGAAGAACAGCGGAGCGCGGCTGCTG 237
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Search completed: April 11, 2003, 23:54:13
 Job time : 188.754 secs

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 10, 2003, 23:54:09 ; Search time 12816.2 Seconds

(without alignments)
5767.767 Million cell updates/sec

Title: US-09-761-640-3

Perfect score: 2540

Sequence: 1 ccgcgtccctcgcgcgtccagcag.....aaaaaaaaaaaaaaaaaaaaa 2540

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenBml :
1: gb_da:*
2: gb_htg:*
3: gb_in:*
4: gb_ov:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rnd:*
36: em_htg_man:*
37: em_htg_vtl:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2151	84.7	2808	9 AK074432	AK074432 Homo sapi
2	2115	83.3	2781	6 AX099939	AX099939 Sequence
3	1930.4	76.0	2905	9 AK094226	AK094226 Homo sapi
4	1641	64.6	1905	9 BC004176	BC004176 Homo sapi
5	1641	64.6	1905	9 BC004210	BC004210 Homo sapi
6	1345.6	53.0	1980	6 AX451343	AX451343 Sequence
7	1303.6	51.3	1755	6 AX398768	AX398768 Sequence
8	1303.6	51.3	1755	6 AK001790	AK001790 Homo sapi
9	1270.6	50.0	2322	6 AX398774	AX398774 Sequence
10	1224.2	48.2	2064	6 AK000522	AK000522 Homo sapi
11	1216.4	47.9	2061	6 AX405828	AX405828 Sequence
12	1028	40.5	1416	6 AX451362	AX451362 Sequence
13	1012	39.8	160903	2 AP002776	AP002776 Homo sapi
14	1012	39.8	171566	2 AP001885	AP001885 Homo sapi
15	1003.4	39.5	1416	9 AB072360	AB072360 Homo sapi
16	797.6	31.4	2736	10 BC028922	BC028922 Mus muscu
17	779.6	30.7	1026	6 AX086034	AX086034 Sequence
18	573.6	22.6	580	9 HUMY153C10	HUMY153C10 Homo sapi
19	558.6	22.0	599	6 AX398772	AX398772 Sequence
20	397	15.6	409	6 AX398771	AX398771 Sequence
21	336	13.2	426	6 AX398765	AX398765 Sequence
22	324.4	12.8	334	6 AX337360	AX337360 Sequence
23	241	9.5	494	6 AX398773	AX398773 Sequence
24	234.2	9.2	279	11 G22634	G22634 human SRS W
25	216.6	8.5	4417	3 AB036834	AB036834 Drosophila
26	212	8.3	1949	6 AX223960	AX223960 Sequence
27	212	8.3	2280	9 AB072356	AB072356 Homo sapi
28	212	8.3	2433	9 AK095421	AK095421 Homo sapi
29	212	8.3	3488	6 AX406972	AX406972 Sequence
30	212	8.3	3817	9 AB072355	AB072355 Homo sapi
31	212	8.3	6374	6 AX180874	AX180874 Sequence
32	208.4	8.2	1052	6 AX223966	AX223966 Sequence
33	208.4	8.2	1711	6 AX223964	AX223964 Sequence
34	201.8	7.9	1796	9 AB072358	AB072358 Homo sapi
35	182.4	7.2	2260	6 AX180876	AX180876 Sequence
36	174.4	6.9	1771	6 AX202239	AX202239 Sequence
37	142.8	5.6	386	9 AF484838	AF484838 Homo sapi
38	142	5.6	113474	2 AC013932	AC013932 Drosophila
39	142	5.6	181132	3 AC008206	AC008206 Drosophila
40	142	5.6	227398	3 AE003750	AE003750 Drosophila
41	121.4	4.8	5463	9 AB037719	AB037719 Homo sapi
42	104.4	4.1	1390	6 AX237020	AX237020 Sequence
43	103.4	4.1	531	6 AX237022	AX237022 Sequence
44	98.8	3.9	1236	6 AX354534	AX354534 Sequence
45	98.8	3.9	1693	9 BC007771	BC007771 Homo sapi

ALIGNMENTS

RESULT 1
AK074432
LOCUS Homo sapiens cDNA FLJ23852 fls, clone KAT12021.
ACCESSION AK074432
VERSION AK074432.1 GI:18677033
KEYWORDS oligo capping; fls (full insert sequence).
SOURCE Homo sapiens signet-ring cell carcinoma cell_line:KATO III cDNA to mRNA, clone_11b:KAT clone:KAT12021.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Watanabe,K., Kunagai,A., Itakura,S., Yamazaki,M., Tashiro,H.,
Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,

Pred. No. is the number of results predicted by chance to have a

TITLE	JOURNAL	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Nakamura, Y., Isogai, T. and Sugano, S.						
NEDO human cDNA sequencing project						
2 (bases 1 to 2808)						
Sugano, S., Suzuki, Y., Ota, T., Obayashi, M., Nishi, T., Isogai, T., Shibahara, T., Tanaka, T. and Nakamura, Y.						
Direct Submission						
Submitted (14-FEB-2002)						
Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan						
(E-mail: shioa@ims.u-tokyo.ac.jp, Tel: 81-3-5449-5286, Fax: 81-3-5449-5416)						
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan: cDNA full insert sequencing: Research Association for Biotechnology: cDNA library construction: 5'- & 3'-end one pass sequencing: Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).						

FEATURES	Location/Qualifiers
source	1. .2808

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="KAT12021"
/cell_line="KATO III"
/cell_type="signed-ring cell carcinoma"
/clone_lib="KAT"
/notes="cloning vector: pMT85FL3"
97..2076
CDS

```

```

/note="unnamed protein product"
/codon_start=1
/odon_start=1
/protein_id="BAB85080.1"
/db_xref="GI:18677034"
/translation="MALVTYSRSPPGSGASTPVGPMDOAVQRSRLOQRFAVLRGH
VLGGDGGDDNDAAEASSEPPEKAPSESEELGGDTDDGGSGSGPQKEQGHILHLM
QLRPDDIRLAOLEAPRPRRLYLIVSTRREEGLSDETVLVGGFPPSSPST
LGLVPLMSDTQYVLDDGGSVTSGGSGSRFKRISLTQMTATLOVLIHQAEALGSL
LVGGASLTAMASHYQERLNSBQSLNEMTAADELSLRPAFSGESQEDMCOAIFL
AELMKYLDVDSLEVSYKREIKQALELRGLGLEQYRDTIDNQMLLVAQRASITFE
HLTIGSEMANAELEIDRNRYTHILNARELEIDNYPERTYTHANLWDEEAOQLPRH
KEHTRLEAARAOQTHVLVHKMGVSNASAALVLANQKVECSLEQALRHVQELRPLR
PNPGFLROLDIYQGIILTASRKHSHWDEKGVSGSEHPAPEVSTPFPFLPEPDEGGGG
EKKVGMDEEAOAPKEBGPSPRINLGKMSILPELSLETSESTSDMPVEFSS
ESSHSEELPEPOLAPTRKGGQVDRGQPOLAKSRSGSVTLTGSAVVRNRTGAPEQED
GGQGGQGEPCITSNPRSRKRVQRA SVHDSPEGGA "

```

BASE COUNT	585 a	912 c	823 g	488 t
ORIGIN				
Query Match	84.7%	Score 2151	DB 9	Length 2808
Best Local Similarity	89.4%	Pred. NO. 0		
Matches 2506	Conservative	0	Mismatches	5
			Indels	292
			Gaps	6

Qy 1 CCTGATCCCTGCGGGGNCAGACTGCCCGGAGGTGTAGAGGAAGGGGGCGTCCCGGTCC 60
 11 CTTGATCCCTGCGGGGTCAGACTGT-CCGCGAGGTTGAGGGAAGGGGGCGTCCCGGTCC 69
 Qy 61 CAGCCAGAGTCTGCGGGGCTGTGCTCAATGAGCCCGTGTACAGTATAGCGGTTCGCCCCG 120

Db 70 CAGCCCAAGTCTCGCGGCTGGCTCCATGGCCCTGGTACAGTACAGCCGTTCCGCCCCG 129

QY 121 GGCAGCGGCGCCCTCCACGCCGCTGGGGCCCTTGGACACAGCGGTCCAGCAGGAAGTCTGA 180

Db 130 GGCAGCGGCGGCTCCACGCCCGTGGGGCCCTTGGACACAGCGGTCCAGCAGGAAGGAGTCA 189

[illegible]

301 GAGGAGCTCCACGGGGACGACGACTTTCGGGGCAAGATCCAGAGTCCCCAGAAGCAG 360

Db	310	GAGGAGCTCCACGGGGACCGAGACGACTTTGGGCGAAGATCCAGATGCCCGCAGAAGCAG	369
QY	361	GAGGAGCAGAGGACGACACCTGCACCTTCATGGTATGATGCTGTGTGAGGGCCGGCAGATGTGACATC	420
	370	GAGGAGCAGAGGGCAGCAGCCTGCAACCTCATGTGTACAGGTGTGAGGGCCGGCAGATGTGACATC	429
QY	421	CGCTGGCAGCCACGACTGAGAGCACCCCGGCTCCCGGCTCCGCTAAGTCTGCTGTGATTT	480
Db	430	CGCTGGCAGCCCGCAGCTGAGAGGACACCCGGGCTCCCGGCTCCGCTAAGTCTGCTGTGATTT	489
QY	481	TCTACAGAGAGAGAGAGGTCTGAGGCAGAGATGAGAGCGGTCTCTCGGCGCGTGATTTTC	540
Db	490	TCTACAGAGAGAGAGAGGTCTGTAGCCAGAGATGAGAGCGGTCTCTCGGCGCGTGATTTTC	549
QY	541	CCTACAGCAGACTCCCCAGCTGCACCTTGAGGCTGTGCTTGCCTTGTGAGATGACACC	600
Db	550	CCTACAGCAGACTCCCCAGCTGCACCTTGAGGCTGTGCTTGCCTTGTGAGATGACACC	609
QY	601	CAGGTACTTACATGAGAGACGGGGGCTTCAGCGTGAAGCTGTGGTGGGCAAAAGCCGGAATC	660
Db	610	CAGGTACTTACATGAGAGACGGGGGCTTCAGCGTGAAGCTGTGGTGGGCAAAAGCCGGAATC	669
QY	661	TTTCAAGCCCATCTCCATCTCAGACCATGT-----	688
Db	670	TTTCAAGCCCATCTCCATCTCAGACCATGTGGGCAACATCCAGGTATTGGACCAAGCATGT	729

QY	689	-----	688
	730	GAGCAGCCTTAGGACAGCGGCTTTGTAACGGGAGTGACGTTGACCTGAGGCGCAGCCAC	789
QY	689	-----	688

Db	790	TACCAGAGAGACTGAACTCCGACAGAGCTGCCCTCATISAGTSGAGCAGGCTATGGCCGAC	849
Qy	689	-----GGTCTCAGACAGGAGCATG	711
Db	850	CTGAGTCTCTGGGGCCTCCGACGGCCGAGACCTGGGGGCTCTCCAAACAGGAGCGAGATG	909

QY	112	GAGCAGGCGGATCCGTCGTCGAGCGTCGTGGAAAGTCGTCGATGTCACGTCACCCGAGGAGATGTTC	777
Db	910	GAGCAGGCGGATCCGTCGTCGATGTGTGGAAAGTCGTGGATGTACGTGACCTGGAGAGATGTTC	969
QY	772	ACTTCCAAAGAGATCGCGCCAGAGGCTCTTGAGCTGCGCTCTGGGGCTCCCCCTCCACAGCATGC	831
Db	970	ACTTCCAAAGAGATCGCGCCAGGCTCTTGAGCTGCGCTGGGGGCTCTCCCTCCACAGCATGC	1024

QY 832 CGTACTTTCATCGACAACCGATGCTGCTCTGTTGGACACGCGGACCGAGCTTCCGC 891
1030 CGTGACTTCATCGACACACGAGATGCTGCTCTGTTGGACACGCGGACCGAGCTTCCGC 1088
QY 892 ATCTCCGCCACCTCTACCTCGGGCTAGAGTGGAGACGCAACCACTTGGAGGAGCTTCGAC 951

Db	Qy	Db	Qy
1090	952	1090	952
ATCTTCCCCCACCCTTACCTTGGGCTCGAGATGTGAACGCAACCAACCTGGAGGAGCTGAC	AGCAACAGGGTCACCCACATCTTGAACATGTGGCCGGGAGATTACACACTTCAACCTGAG	ATCTTCCCCCACCCTTACCTTGGGCTCGAGATGTGAACGCAACCAACCTGGAGGAGCTGAC	AGCAACAGGGTCACCCACATCTTGAACATGTGGCCGGGAGATTACACACTTCAACCTGAG
114	101	114	101
1150	1150	1150	1150
AGCAACAGGGTCACCCACATCTTGAACATGTGGCCGGGAGATTACACACTTCAACCTGAG	AGCAACAGGGTCACCCACATCTTGAACATGTGGCCGGGAGATTACACACTTCAACCTGAG	AGCAACAGGGTCACCCACATCTTGAACATGTGGCCGGGAGATTACACACTTCAACCTGAG	AGCAACAGGGTCACCCACATCTTGAACATGTGGCCGGGAGATTACACACTTCAACCTGAG
120	120	120	120

QY	1012	CGCACCCATTCATGCGCCTGGAGTAGAGAGTGGCCAGCGCGCCGCAC	107
QY <td>1012</td> <td>CGCACCCATTCATGCGCCTGGAGTAGAGAGTGGCCAGCGCGCCGCAC</td> <td>107</td>	1012	CGCACCCATTCATGCGCCTGGAGTAGAGAGTGGCCAGCGCGCCGCAC	107
Db	1210	CGCTTACCCATTCATGCGCCTGGAGTAGAGAGTGGCCAGCGCGCCGCAC	126
QY	1072	TGGAGAGAGCGCACCGCTTCATTTGAGCGCTGCAAGAGCAGAGGGCACCCACGCGTGGTGC	113
Db	1270	TGGAGAGAGCGCACCGCTTCATTTGAGCGCTGCAAGAGCAGAGGGCACCCACGCGTGGTGC	1320

D0	1GGAAAGAGAGAGACACCCCTTCATTCATTCGAGGCTGCAAGAGACATAGGGCCACCCACACGTCGTCGATC	133
D1	12470	
Qy	1132 CACGCGAAGATGGGGCGTCGAGCCGCTCTAGCGGGCCACAGTGGCTGGCCATATGCCATGAAGCAG	119
Db	1330 CACTGCAGATGGGGCGTCAGACCCGCTCTAGCGGGCCACAGTGGCTGGCCATATGCCATGAAGCAG	138
Y	1192 TAGCAATGACGCTTGAGCAGGGCCCTGCGGCACGTCAGAGAGTCCGGGCCATATCGGCCCG	125

Db 1390 TACGAATGCAgCCTGGAGCAGGCCCTGCGCCACGTGAGAGGCTCCGGCCATGCGCCGC 144

QY	1252	CCCAACCTGGCTTCGCGGCGACGCTGACATCTACAGGGCATCTACG-----	1302
Db	1450	CCCAACCTGGCTTCGCGGCGACGCTGACATCTACAGGGCATCTTACGCGCACGCC	1509
QY	1303	-----	1302
Db	1510	CAGAGCCATGTCTGGGAGCGAAAGTGGGTGGGTCTCCCGCAGAGAGACACCAGCCCT	1569
QY	1303	-----GCCAAGACCTGAGGGGTGTGGGAGAGAG	1333
Db	1570	GAAGTCTCTACACCATTCACACCTCTTCGCGCCAGAAACCTGAGGGGTGTGGGAGAGAG	1629
QY	1334	GTTTTGTAGCATGGAAGAGAGACGAGCGACGCCCGAAAGAGACCTGTGGGCGCACGGGAC	1393
Db	1630	GTGTGTAGCATGGAAGAGAGACCGAGCGACGCCCGAAAGAGAACCTGGG--CCAGCGCAC	1687
QY	1394	GTTAAACCTTCGAGGGGTGATGAGGTTCATCATCATGTCTTGTGAGCCCTCTTGGGAGCTG	1453
Db	1688	GTTAAACCTTCGAGGGGTGATGAGGTTCATCATCATGTCTTGTGAGCCCTCTT--GGAGCTG	1746
QY	1454	GAGAGCACTCTAGAGACCGATGACATGTGCAGAGGGCTCTCTTCCGACGAGTCTTCA	1513
Db	1747	GAGAGCACTCTAG--AGACAGTATGACATGCGAGAGGGCTTCTTCTTCCAGAGTCTTCA	1805
QY	1514	TGAAGACCTCTGCGAGCCCTTTCACAGCTTGTGCAAGAGCAAGGAGGCGCAGAGTGTGA	1573
Db	1806	TGAAGACCTCTGCGAGCCCTTTCACAGCTTGTGCAAGAGCAAGGAGGCGCAGAGTGTGA	1855
QY	1574	CAGGGGCGCTCAGCCCTGCTGTAAGTCCCGCCAGTCAGTGTGTTACCTCTCCAGGCGATGC	1633
Db	1866	CAGGGGCGCTCAGCCCTGCTGTAAGTCCCGCCAGTCAGTGTGTTACCTCTCCAGGCGATGC	1925
QY	1634	CGTGTGGCCCAACCGGACCGAGGCGCTTCCAGAGAGCGAGAGCGAGGGGCGAGGGGCA	1693
Db	1926	CGTGTGGCCCAACCGGACCGAGGCGCTTCCAGAGAGCGAGAGCGAGGGGCGAGGGGCA	1985
QY	1694	GGGAGAGCCCTGCATTTCTCTACGCCCAAGTTCCGGAAGTGTGTGAGACAGGCCAGCT	1733
Db	1986	GGGAGAGCCCTGCATTTCTCTACGCCCAAGTTCCGGAAGTGTGTGAGACAGGCCAGCT	2045
QY	1754	GCAATGAGTGTGAGAGAGGGGCGAGGCGCTGAGCGCTTCACATATGCCACGCTCCCTAC	1813
Db	2046	GCAATGAGTGTGAGAGAGGGGCGAGGCGCTGAGCGCTTCACATATGCCACGCTCCCTAC	2105
QY	1814	ACTGAGAGGATCCACACTCTCTTGGAGAAACCCCTCAGCTGTGTGCCGACACATTC	1873
Db	2106	ACTGAGAGGATCCACACTCTTGTGAGAAACCCCTCAGCTGTGTGCCGACACATTC	2165
QY	1874	CTCTCAGCTCTCGGCCCATACCCTGTACTAGCGTCACTCCACCTCCAGCCCTGTACTAGCGC	1933
Db	2166	CTCTCAGCTCTCGGCCCATACCCTGTACTAGCGTCACTCCACCTCCAGCCCTGTACTAGCGC	2225
QY	1934	CTCAGCTCCACCCCTGTGCTACTACAGCTCACTCAGCTTACAGCTTAACTCCAGGCCAT	1993
Db	2226	CTCAGCTCCACCCCTGTGCTACTACAGCTCACTCAGCTTACAGCTTAACTCCAGGCCAT	2285
QY	1994	GTCCTGCTGTCCAAAGGGCTTAAGACTTTCTAACTGGGATGTGTGTAGAGGACTGAAGTA	2053
Db	2286	GTCCTGCTGTCCAAAGGGCTTAAGACTTTCTAACTGGGATGTGTGTAGAGGACTGAAGTA	2345
QY	2054	CTTTTGGGGCAACAGACCCCTAGTTCACTTCAACTGTACCTCTGAGACACTCACTCGT	2113
Db	2346	CTTTTGGGGCAACAGACCCCTAGTTCACTTCAACTGTACCTCTGAGACACTCACTCGT	2405
QY	2114	GGCAGCGAATGAAACAGAGCTTCCGCTGCAAAAAGGTCACGCTCCACCCCGGCC	2173
Db	2406	GGCAGCGAATGAAACAGAGCTTCCGCTGCAAAAAGGTCACGCTCCACCCCGGCC	2465
QY	2174	CTTCCCTCAGCTCTGTGCTCTCCCAAGTTCACTTCCGGAACCGCAGGCCAGGCAACCA	2233
Db	2466	CTTCCCTCAGCTCTGTGCTCTCCCAAGTTCACTTCCGGAACCGCAGGCCAGGCAACCA	2535

QY	2234	GTGCCCCCAAGGACGAGATGCTCAGAGCCGACGCGGAGAGGCTGGAAGGGCTGG	2293
Db	2526	GTGGGCCCCAAAGGACGAGGATGCTCTAGGCCCCAGCGCGGGAGGCTGGGAAGGGCTGG	2585
QY	2234	CAGATCGGCTTCCTCATCCACTCCACCGGATCCAGGCTCTTGCTCTGTCCCAACCTC	2353
Db	2586	CAGATCGGCTTCCTCATCCACTCCACCGGATCCAGGCTCTTGCTCTGTCCCAACCTC	2645
QY	2354	CTGTACACACACGCCAGATCACAGGGGACACAGGCCAGAGATAGTCTTCTTTTGTCCTT	2413
Db	2646	CTGTACACACACGCCAGATCACAGGGGACACAGGCCAGAGATAGTCTTCTTTTGTCCTT	2705
QY	2444	CTGGCCTCTGGCTAGTCAGTTTTCATATAGCCTTACAGTATCTGCGCTTGTGATGGAAT	2473
Db	2706	CTGGCCTCTGGCTAGTCAGTTTTCATATAGCCTTACAGTATCTGCGCTTGTGATGGAAT	2765
QY	2474	AAACACATTTTCATATTTGGTTAAAAAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	2516
Db	2766	AAACACATTTTCATATTTGGTTAAAAAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	2808

LOCUS	AX099939	2781 bp	DNA	linear	PAT 02-APR-2001
DEFINITION	Sequence 21 from Patent WO0120004.				
ACCESSION	AX099939				
VERSION	AX099939.1	GI:13538949			
KEYWORDS					
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 2781)				
AUTHORS	Yue, H., Tang, Y.T., Bandman, O., Hillman, J.L., Baughn, M.R., Azimzadeh, Y. and Lu, D.A.				
TITLE	Protein phosphatase and kinase proteins				
JOURNAL	Patent: WO 0120004-A 21 22-MAR-2001; Incyte Genomics, Inc. (US)				
FEATURES	Location/Qualifiers				
source	1..2781				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/note="Incyte ID NO: 5039718CB1"				
BASE COUNT	576 a 906 c 820 g 479 t				
ORIGIN					
Query Match	83.3%; Score 2115; DB 6; Length 2781;				
Best Local Similarity	89.3%; Pred. NO. 0;				
Matches 2470; Conservative	0; Mismatches 5; Indels 292; Gaps				
QY	10 GCGGGTCCAGAGACTGTCCCGCGGGGTTGAGGGAAGGGCCGTGCCGTGCCAGCCAGG	69			
DB	10 GCGGGTCCAGAGACTGT -CGCGCGGTTGAGGGAAGGGCCGTGCCGTGCCAGCCAGG	68			
QY	70 TGCCTGCGGCGCGGTGCTCATGGCCCTGTGTCACAGTACGCCCTTCCGCCCGCGGACGCGG	129			
DB	69 TGTCTGCGGCGCGGTGCTCATGGCCCTGTGTCACAGTACGCCCTTCCGCCCGCGGACGCGG	128			
QY	130 GCCTCCAGCGCCCGTGTGGGGCCCTGGAGACCAAGGGGATGTCACAGCGAAGATGTCAGCTCCAGCGA	189			
DB	129 GCCTCCAGCGCCCGTGTGGGGCCCTGGAGACCAAGGGGATGTCACAGCGAAGATGTCAGCTCCAGCGA	188			
QY	190 AGGCAGAGCTTTGGCGGTGCTCCGTGGGGCTGTCTTGGGACTGACAGATGAGAGGGACAAAT	249			
DB	189 AGGCAGAGCTTTGGCGGTGCTCCGTGGGGCTGTCTTGGGACTGACAGATGAGAGGGACAAAT	248			
QY	250 GATGATGAGAGCAGAGAGCCAGTTCTGAGCCAACAGAGAAGGCCCGAGTGGAGAGGAGGCTC	309			
DB	249 GATGATGAGAGCAGAGAGCCAGTTCTGAGCCAACAGAGAAGGCCCGAGTGGAGAGGAGGAGCTC	308			
QY	310 CACGGGGACACACAGACTTGGGGCAAGATCCAGAGTCCCTCCAGAAACAGAGAGGACAG	369			
DB	309 CACGGGGACACACAGACTTGGGGCAAGATCCAGAGTCCCTCCAGAAACAGAGAGGAGGACAG	368			

QY 370 AGGACACCTGACCTCATGTGTACAGCTGGTGAAGGCCGACAGATGACATCCGCTGGCA 429
 Db 369 AGGACGACCTGACCTCATGTGTACAGCTGGTGAAGGCCGACAGATGACATCCGCTGGCA 428
 QY 430 GCCCAGCTGAGAGCAGCCCGGCTCCCGGCTCCGCTACCTGCTGTAGTTCCTACAGCA 489
 Db 429 GCCCAGCTGAGAGCAGCCCGGCTCCCGGCTCCGCTACCTGCTGTAGTTCCTACAGCA 488
 QY 490 GAAGGAGAGGCTGAGGACAGATGAGAGCGTCTGCGGGGTGGATTTCCCTACAGC 549
 Db 489 GAAGGAGAGGCTGAGGACAGATGAGAGCGTCTGCGGGGTGGATTTCCCTACAGC 548
 QY 550 AGCTCCCGCAGCTGACCTGCGGCTGGTCTTGGCCCTCTGAGTGAACCCAGGCTGTAC 609
 Db 549 AGCTCCCGCAGCTGACCTGCGGCTGGTCTTGGCCCTCTGAGTGAACCCAGGCTGTAC 608
 QY 610 TTAGATGAGAGCGGGGCTTACAGCTGACGTCTGGTGGGCAAGCCGATCTTCAAGCC 669
 Db 609 TTAGATGAGAGCGGGGCTTACAGCTGACGTCTGGTGGGCAAGCCGATCTTCAAGCC 668
 QY 670 ATCTCATCCAGACCATGT----- 688
 Db 669 ATCTCATCCAGACCATGTGGGCCACACTCAGGTTATTGCCAAGCATGTGAGCAGCT 728
 QY 689 ----- 688
 Db 729 CTAGGACAGCGGCTTGTACCGGGTGGAGTGCCTCACCTGGGGCAGGCATACAGGAG 788
 QY 689 ----- 688
 Db 789 AGACTGAACCTCCGAACAGAGCTCCTCATGATGAGTGGAGCGGTATGCGCAGCTGAGTCT 848
 QY 689 -----GGTCTCAGAACAGAGCAGATGAGAGCGG 720
 Db 849 CTGCGGCTCCAGTGGCCGAGCTGGCGGTCTCTAGAACAGAGCAGATGAGAGCGG 908
 QY 721 ATCCGTGCTGAGCTGTGAAAGTGTGATGTCACTGACCTGAGAGTGTCACTTCCAA 780
 Db 909 ATCCGTGCTGAGCTGTGAAAGTGTGATGTCACTGACCTGAGAGTGTCACTTCCAA 968
 QY 781 GAGATCCCGCAGGCTGTGAGTGGCGCTGGGGCTCCCTCCACAGTACCTGACTTC 840
 Db 969 GAGATCCCGCAGGCTGTGAGTGGCGCTGGGGCTCCCTCCACAGTACCTGACTTC 1028
 QY 841 ATGCACACAGATGCTGCTGTGGTGGACAGCGGACCGGCTCCCGCATCTTCCCC 900
 Db 1029 ATGCACACAGATGCTGCTGTGGTGGACAGCGGACCGGCTCCCGCATCTTCCCC 1088
 QY 901 CACCTCTACCTGGGCTCAGAGTGAAGCGACAACTGGAGGAGCTGCAGAGAAACAG 960
 Db 1089 CACCTCTACCTGGGCTCAGAGTGAAGCGACAACTGGAGGAGCTGCAGAGAAACAG 1148
 QY 961 GTACCCACATTTGAACATGGCGCGGAGATTGACAACTTACCTGAGGCTTCAAC 1020
 Db 1149 GTACCCACATTTGAACATGGCGCGGAGATTGACAACTTACCTGAGGCTTCAAC 1208
 QY 1021 TACCAAAATGTGCGCTGTGGAGTGAAGTGGCCAGCTGCTCCGCTGAGGAAGAG 1080
 Db 1209 TACCAAAATGTGCGCTGTGGAGTGAAGTGGCCAGCTGCTCCGCTGAGGAAGAG 1268
 QY 1081 ACGCACCCTTCATTGAGGCTGCAAGAGCAGAGGCGACCCAGCTGCTGCTCACTGCAAG 1140
 Db 1269 ACGCACCCTTCATTGAGGCTGCAAGAGCAGAGGCGACCCAGCTGCTGCTCACTGCAAG 1328
 QY 1141 ATGGGCGTACGCGCTGACAGCGGACACAGTGTGGCTATGCCATGAAGATGCAATG 1200
 Db 1329 ATGGGCGTACGCGCTGACAGCGGACACAGTGTGGCTATGCCATGAAGATGCAATG 1388
 QY 1201 AGCTGAGACAGGCGCTGGCGGACAGTGTCCGGCCATGCGCGCCGACCAACCT 1260
 Db 1389 AGCTGAGACAGGCGCTGGCGGACAGTGTCCGGCCATGCGCGCCGACCAACCT 1448

QY 1261 GGCCTTCTGCGCCAGCTGACATCTACAGGGGATCCTGAGC----- 1302
 Db 1449 GGCCTTCTGCGCCAGCTGACATCTACAGGGGATCCTGAGCGCCAGCGCAGAGCAT 1508
 QY 1303 ----- 1302
 Db 1509 GTCTGGAGACAGAAAGTGGGTGCTCCAGAGAGCACCCAGCCCTGAAGTCT 1568
 QY 1303 -----GCCAAGCTGAGGCTGGTGGGAGAGAAAGTTGTAGC 1342
 Db 1569 ACACCATTTCCACTCTTCCGACAGAACTGAGGCTGGTGGGAGAGAAAGTTGTAGC 1628
 QY 1343 ATGAAGAGAGCAGAGCAGCCCGGAAAGAGCCTGGGCGCACGGGACGATATAAC 1402
 Db 1629 ATGAAGAGAGCAGAGCAGCCCGGAAAGAGCCTGGG--CCAGCGCACGATATAAC 1686
 QY 1403 TCCGAGGGGTCTATGAGTTCATCACTCTTGTGAGCCCTCTCTGGGAGCTGGAAGAC 1462
 Db 1687 TCCGAGGGGTCTATGAGTTCATCACTCTTGTGAGCCCTCTCT--GGAGCTGGAAGAC 1745
 QY 1463 TCAGTAGACAGTGAATGACATGCGCAGAGGCTCTCTTCCACAGAGTCTTCAATGAAG 1522
 Db 1746 TCAG--AGACAGTGAATGACATGCGCAGAGGCTCTCTTCCACAGAGTCTTCAATGAAG 1804
 QY 1523 TCTGACGCCCTTCCACAGCTTGAAGGACAGGAGGCGCAGAGTGAAGGCGC 1582
 Db 1805 TCTGACGCCCTTCCACAGCTTGAAGGACAGGAGGCGCAGAGTGAAGGCGC 1864
 QY 1583 TCAGGCTGCTTGAAGTCCCGGACATCACTGCTTACCTCCAGGGCAGTCCGTGGTGG 1642
 Db 1865 TCAGGCTGCTTGAAGTCCCGGACATCACTGCTTACCTCCAGGGCAGTCCGTGGTGG 1924
 QY 1643 CAACGAGACCCAGGCTTCCAGAGCAGAGGAGGCGCAGGGGACAGGAGAGC 1702
 Db 1925 CAACGAGACCCAGGCTTCCAGAGCAGAGGAGGCGCAGGGGACAGGAGAGC 1984
 QY 1703 CTGCAATTTCTTCTACGCCAGGCTCCGGAAGTGTGAGACAGCCAGCTGATGACAG 1762
 Db 1985 CTGCAATTTCTTCTACGCCAGGCTCCGGAAGTGTGAGACAGCCAGCTGATGACAG 2044
 QY 1763 TGGAGAGGAGGCGGAGGCTGAGGCTCAGACATGCGCAGGCTCCCTGACACTGAAG 1822
 Db 2045 TGGAGAGGAGGCGGAGGCTGAGGCTCAGACATGCGCAGGCTCCCTGACACTGAAG 2104
 QY 1823 GATCCACAATCTCTTGAAGAAACACCTTCACTGCTGTTGGCCGACACATCTCTCAG 1882
 Db 2105 GATCCACAATCTCTTGAAGAAACACCTTCACTGCTGTTGGCCGACACATCTCTCAG 2164
 QY 1883 CCGGCCATACCGCTTACTACAGCTCACCTCCACACCTGCTACACTAGGCTCACCTCC 1942
 Db 2165 CCGGCCATACCGCTTACTACAGCTCACCTCCACACCTGCTACACTAGGCTCACCTCC 2224
 QY 1943 CACCCCTGTCACTACAGCTCACCTTCACTGCTTAACTGAGGAGGCTGAGGCTGAG 2002
 Db 2225 CACCCCTGTCACTACAGCTCACCTTCACTGCTTAACTGAGGAGGCTGAGGCTGAG 2284
 QY 2003 TCCAGGGGCTCAAGATTTCTACTGAGTGTGATGAGAGGACTGAAGTACTTTGGGG 2062
 Db 2285 TCCAGGGGCTCAAGATTTCTACTGAGTGTGATGAGAGGACTGAAGTACTTTGGGG 2344
 QY 2063 GCAACAGACACCTAGTTTCTATTTCTCACTTACCTGTGACACTGCTGTGACAGGAA 2122
 Db 2345 GCAACAGACACCTAGTTTCTATTTCTCACTTACCTGTGACACTGCTGTGACAGGAA 2404
 QY 2123 TGAAGAGAGGCTCCCGGCAAGAGGCTGAGGCTCCAGCCCGGCGCTCTCTGCA 2182
 Db 2405 TGAAGAGAGGCTCCCGGCAAGAGGCTGAGGCTCCAGCCCGGCGCTCTCTGCA 2464
 QY 2183 CCTCTGTCTCTCCAGTTCTTCTGGAACACAGGCGCAGGCAACCACTGAGCCCC 2242
 Db 2465 CCTCTGTCTCTCCAGTTCTTCTGGAACACAGGCGCAGGCAACCACTGAGCCCC 2524
 QY 2243 AAAGGAGGACAGATCTCAGGCGCCAGCGGAGGCTGGAAGGCTGGCAGATCGCT 2302

|||||
Db 2525 AAGGCGAGGAGATCCTCAGGCGCCGAGCGGAGGCTGGAAGGCGCTGCAGATCGCT 2584
Qy 2303 TCCCTCATCCACCCCGCCGCTCCAGCTCTTGTCTGCTGCCAGACCTCTGTGACAC 2352
Db 2585 TCCCTCATCCACCCCGCCGCTCCAGCTCTTGTCTGCTGCCAGACCTCTGTGACAC 2644
Qy 2363 CAGCCGAGATCACAGGCGCAGGCGCAGATAGCTCTTGTCTGCTCTGTGCGCTCT 2422
Db 2645 CAGCCGAGATCACAGGCGCAGGCGCAGATAGCTCTTGTCTGCTCTGTGCGCTCT 2704
Qy 2423 GCGTAGTCAGTTTTCATAGCCTTACAGATCTGCTGCTTGTGATGAGAATAAACAAT 2482
Db 2705 GCGTAGTCAGTTTTCATAGCCTTACAGATCTGCTGCTTGTGATGAGAATAAACAAT 2764
Qy 2483 TTTTCATA 2489
Db 2765 TTTTCATA 2771
RESULT 3
LOCUS AK094226 2905 bp mRNA linear PRI 15-JUL-2002
DEFINITION Homo sapiens cDNA FLJ36907 fis, clone BRACE2003800, weakly similar
to MAP kinase phosphatase.
ACCESSION AK094226
VERSION AK094226.1 GI:21753246
KEYWORDS oligo capping, fis (full insert sequence).
SOURCE Homo sapiens cerebellum cDNA to mRNA, clone_11b:BRACE2
clone:BRACE2003800.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Tanigami,A., Fujiwara,T., Shibahara,T., Goto,Y., Hiraio,M.,
Shimizu,F., Wakebe,H., Ono,T., Hishigaki,H., Watanabe,T., Ozaki,K.,
Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S.,
Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T.,
Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M.,
Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K.,
Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y.,
Sugano,S., Nagahara,K., Masuno,Y., Nagai,K. and Isogai,T.
Unpublished
NEDO human cDNA sequencing project
TITLE Unpublished
JOURNAL 2 (bases 1 to 2905)
REFERENCE Isogai,T. and Yamamoto,J.
AUTHORS Direct Submission
TITLE Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team): 2-6-7
JOURNAL Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
REFERENCE (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
AUTHORS NEDO human cDNA sequencing project supported by Ministry of
JOURNAL Economy, Trade and Industry of Japan; cDNA full insert sequencing:
COMMENT Research Association for Biotechnology (RAB): cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation: clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
FEATURES
source
1. .2905
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="BRACE2003800"
/tissue.type="cerebellum"
/clone_11b="BRACE2"
/note="cloning vector: pME18SFL3"
646..2187
/note="unnamed protein product"
/codon.start=1
/protein_id="BAC04314.1"
/db_xref="GI:21753247"
/translation="MAPKIPASSPSCCTGLVPLPLMSDPTQVYLDGDSGVTSYGGGSR
IFPISIQTMATLQVLRHQCEALISGLVPGSALTWASHYQERLNSDSCINEMTA

MADLESLRPPSPAGSGSSSEOEOMEQAI RAEMLRVLDVSDLESYTSKEIRQALERLGL
PLOQYRDEIDNOMLVAORRASRIEPIHLYLSENNAAJLELORNRTHILNMARE
IDNYPREFYHNRILMDESSAOLLPWKETHRFIFARROGTHVLVHCMMYRSRA
TVLAVAMKOYECLEALRIVOLERLAPRNPFLRQLOLYGOLITRASRSHVBEOKV
GGVSPBEPAPVEVSTPPPLPEBEGGKVVGMESCAAPAEERPRIRLVGM
RSLSLEPSELESTSETSDMPVFSHESHSHEPLOPFLARTRGGQVDGDPDA
LKSROSVYVTLQGSVAVVANRTOAFQEOEGOGGOGGCPICISSTPRFRKVVROASVHDS
EEGEA"
BASE COUNT 591 a 940 c 867 g 507 t
ORIGIN
Query Match 76.0%; Score 1930.4; DB 9; Length 2905;
Best Local Similarity 84.5%; Pred. No. 0;
Matches 2457; Conservative 0; Mismatches 6; Indels 443; Gaps 8;
Qy 27 CCGGGGGTTGAGGGAAGGGCCGTGCCGGTGCACGCCACAGTGTCTGCGGCTGCTC 86
Db 3 CCGGGGGTTGAGGGAAGGGCCGTGCCGGTGCACGCCACAGTGTCTGCGGCTGCTC 62
Qy 87 CATGGCCCTGTGACAGTGAAGCCGTTGCCCGGGGAGCGGCGCTTCACGCCCGTGGG 146
Db 63 CATGGCCCTGTGACAGTGAAGCCGTTGCCCGGGGAGCGGCGCTTCACGCCCGTGGG 122
Qy 147 GCCCTG----- 152
Db 123 GCCCTGGAATTCGTGAGAGAGGGGAGGAGCGCTCCCGCCCTCACCCTGCCCTGGG 182
Qy 153 -----GACACAGCGGCGTCCAGCGAAGAGTGTGACTCCAGCGAGAGAG 197
Db 183 GCTCTCTCTGCGCAGGACCGAGGCGTCCAGCGAAGAGTGTGACTCCAGCGA 238
Qy 198 CTTGCGGCTGCTCGTGGGCGTCTGCGGAGTGTGAGTGTGAGGAGGAGTGTGATGC 257
Db 239 TTTTGGGCTGCTCGTGGGCGTCTGCGGAGTGTGAGTGTGAGGAGGAGTGTGATGC 298
Qy 258 AGCAGAGCCAGTTCTGAGCCACAGAAAGAGCCCGAGTGTGAGGAGAGTGTGAGGAG 317
Db 299 AGCAGAGCCAGTTCTGAGCCACAGAAAGAGCCCGAGTGTGAGGAGAGTGTGAGGAG 358
Qy 318 CCAACAGACTTGGGCGAAGATCCCGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 377
Db 359 CCAACAGACTTGGGCGAAGATCCCGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 418
Qy 378 CCTGACCTCATGTGTACAGCTGTGAGCGCCAGAGATGATCCGCTGCGAGCCAGCT 437
Db 419 CCTGACCTCATGTGTACAGCTGTGAGCGCCAGAGATGATCCGCTGCGAGCCAGCT 478
Qy 438 GGAGGACCCCGGCTCCCGGCTCCGCTACCTGCTGTGATTTCTACAGAGAAGAGA 497
Db 479 GGAGGACCCCGGCTCCCGGCTCCGCTACCTGCTGTGATTTCTACAGAGAAGAGA 538
Qy 498 AGGTCTGAGCCAGATGAGAGGCTCTCTGCGGCTGTGATTTCCCTGACAG----- 548
Db 539 AGGTCTGAGCCAGATGAGAGGCTCTCTGCGGCTGTGATTTCCCTGACAGAGTTGCA 598
Qy 549 ----- 548
Db 599 GCAGGAGAGGAAGAGGGGCAAGAGTGTGAGGAGGAGGGGAGCAATGCTTTCCCTC 658
Qy 549 -----CAGTCCCGCCAGCTGCAAGCCCTGGGCTGTGCTTCCCTGTGAGTGA 598
Db 659 TGTCCCTGCAAGCTCCCGCAGCTGCAAGCCCTGGGCTGTGCTTCCCTGTGAGTGA 718
Qy 599 CCCAGGTGTACTTGTAGATGAGAGCGGGGCTTACAGCTGTGCTGTGAGAGAGAGAG 658
Db 719 CCCAGGTGTACTTGTAGATGAGAGCGGGGCTTACAGCTGTGCTGTGAGAGAGAGAG 778
Qy 659 TCTTCAAGCCATCTCATCCAGACATGTGGGCGACAGTGTGATGTGACCAAGAT 688
Db 779 TCTTCAAGCCATCTCATCCAGACATGTGGGCGACAGTGTGATGTGACCAAGAT 838
Qy 689 ----- 688

Dh 839 GTGAGCAGCTTAGGAGCGGCTTGTACGGGTGGCAGTGGCCCTCAGCTGGGCCAGCC 898
QY 689 ----- 688
Dh 899 ACTACAGAGAGACTGAAGTCCGAAACAGAGCTGCCATGAGTGGACGGCTATGGCCG 958
QY 689 -----GTCCTCCAGAACAGAGAGA 709
Dh 959 ACCTGAGTCTCTGGCGGCTCCAGCGCCGAGCCTGGCGGGTCTCTGAAACAGAGAGAGA 1018
QY 710 TGGAGCAGGCGATCCGCTGCTGAGCTGTGGAAGTGTGGATGTGAGTACCTGAGAGAGTG 769
Dh 1019 TGGAGCAGGCGATCCGCTGCTGAGCTGTGGAAGTGTGGATGTGAGTACCTGAGAGAGTG 1078
QY 770 TCACCTTCAAAGAGATCCGCGAGCTGTGAGTGGCGCCGGGGGCTCCCTCCAGCAGT 829
Dh 1079 TCACCTTCAAAGAGATCCGCGAGCTGTGAGTGGCGCCGGGGGCTCCCTCCAGCAGT 1138
QY 830 ACCGTGACTTCATGCAACAAACAGATGCTGCTGTGGTGGCACAGCGGAGACGAGCTGCC 889
Dh 1139 ACCGTGACTTCATGCAACAAACAGATGCTGCTGTGGTGGCACAGCGGAGACGAGCTGCC 1198
QY 890 GCATCTTCCCGCCTCTACCTGCGGCTGCAAGTGGAAACGACGAAACCTGAGAGAGCTGC 949
Dh 1199 GCATCTTCCCGCCTCTACCTGCGGCTGCAAGTGGAAACGACGAAACCTGAGAGAGCTGC 1258
QY 950 AGAGAACAGGGGTACCCACATCTTGAACATGCGCCGGGAGATTGACACTTCTACCTG 1009
Dh 1259 AGAGAACAGGGGTACCCACATCTTGAACATGCGCCGGGAGATTGACACTTCTACCTG 1318
QY 1010 AGCGCTTCACCTACCAACATGTGGCGCTGTGGATGAGAGTGGCCAGCTGTGCTGCGC 1069
Dh 1319 AGCGCTTCACCTACCAACATGTGGCGCTGTGGATGAGAGTGGCCAGCTGTGCTGCGC 1378
QY 1070 ACTGGAAGAGACGACCGCTTCATTTGAGCTGCAAGAGACAGGCGACCCAGCTGCTGG 1129
Dh 1379 ACTGGAAGAGACGACCGCTTCATTTGAGCTGCAAGAGACAGGCGACCCAGCTGCTGG 1438
QY 1130 TCCACTGCAAGATGGGGGCTGACCGCTCAGCGGCCACAGTGTGCTGCTATGCTCATGAAC 1189
Dh 1439 TCCACTGCAAGATGGGGGCTGACCGCTCAGCGGCCACAGTGTGCTGCTATGCTCATGAAC 1498
QY 1190 AGTAGAATGACGCTGAGCAGCGCTGCGGCACGTGAGAGAGTCCGGGCCATCGGCC 1249
Dh 1499 AGTAGAATGACGCTGAGCAGCGCTGCGGCACGTGAGAGAGTCCGGGCCATCGGCC 1558
QY 1250 GCGCCAAACCTTGGCTTCTGCGGCCAGCTGACAGATCTACACAGGGCATCTCTGAAGC 1302
Dh 1559 GCGCCAAACCTTGGCTTCTGCGGCCAGCTGACAGATCTACACAGGGCATCTCTGAAGC 1618
QY 1303 ----- 1302
Dh 1619 GCCAGAGCATGTCTGGAGAGCAAGAAATGGGTGGGGTCTCCCGAGAGAGACCCAGCC 1678
QY 1303 -----GCCAGAACTGAGGGTGGTGGGAGAGA 1331
Dh 1679 CTGAAGTCTTACACCATTTCCACACTTCTCCGCCAAGAACTGAGGGTGGTGGGAGAGA 1738
QY 1332 AGGTGTGAGATGGAAG 1391
Dh 1739 AGGTGTGAGATGGAAG 1796
QY 1392 ACGTATAAACCTCCGAGGGGTCATGAGTGCATGAGTCTTGTGAGAGAGAGAGAGAGAGAG 1451
Dh 1797 ACGTATAAACCTCCGAGGGGTCATGAGTGCATGAGTCTTGTGAGAGAGAGAGAGAGAG 1855
QY 1452 TGGAGAGACCTCAATGAGACAGTGCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1511
Dh 1856 TGGAGAGACCTCAATGAGACAGTGCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1914
QY 1512 CATGAAGAGGCTGTGAG 1571
Dh 1915 CATGAAGAGGCTGTGAG 1974

QY 1572 GACAGGGGGCTCAGCCTGCTGAGATCCCGGACAGTCAAGTGGTTACCTCCAGGGCAGT 1631
Dh 1975 GACAGGGGGCTCAGCCTGCTGAGATCCCGGACAGTCAAGTGGTTACCTCCAGGGCAGT 2034
QY 1632 GCCGTGTGGCCAAACCGGAG 1691
Dh 2035 GCCGTGTGGCCAAACCGGAG 2094
QY 1692 CAGGAG 1751
Dh 2095 CAGGAG 2154
QY 1752 GTGATGACAGTGGAG 1811
Dh 2155 GTGATGACAGTGGAG 2214
QY 1812 ACAGTGAAGAGAGATCCAGACCTCTTGGAGAAACACCTCAGCTGTGTTGCCGACAGAT 1871
Dh 2215 ACAGTGAAGAGAGATCCAGACCTCTTGGAGAAACACCTCAGCTGTGTTGCCGACAGAT 2274
QY 1872 TCCCTCAGCTCCGCGCCATACCTCCGCTACAGCTCAGCTCCAGCCCTGTGCTACG 1931
Dh 2275 TCCCTCAGCTCCGCGCCATACCTCCGCTACAGCTCAGCTCCAGCCCTGTGCTACG 2334
QY 1932 GCGTCACTCCGCGCGCTGCTACAGCTCAGCTCCTTACAGCTTAACTGAGAGAGAG 1991
Dh 2335 GCGTCACTCCGCGCGCTGCTACAGCTCAGCTCCTTACAGCTTAACTGAGAGAGAGAG 2394
QY 1992 ATGTCTGCTGTGCAAG 2051
Dh 2395 ATGTCTGCTGTGCAAG 2454
QY 2052 TACCTTTGGGGGCAACAGACCTTGAATTCATCTCACTGAGAGAGAGAGAGAGAGAGAG 2111
Dh 2455 TACCTTTGGGGGCAACAGACCTTGAATTCATCTCACTGAGAGAGAGAGAGAGAGAGAG 2514
QY 2112 GTGGCAG 2171
Dh 2515 GTGGCAG 2574
QY 2172 CCGTCCCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2231
Dh 2575 CCGTCCCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2634
QY 2232 CAGTGGGCCCCAAG 2291
Dh 2635 CAGTGGGCCCCAAG 2694
QY 2292 GGCAGATGCTTCCCTCATGCACTCCAGCGGTCAGAGTCTTGTGCTGCTGCTGCTGCTGCT 2351
Dh 2695 GGCAGATGCTTCCCTCATGCACTCCAGCGGTCAGAGTCTTGTGCTGCTGCTGCTGCTGCT 2754
QY 2352 TCCGTGACACACAGCGCAGATCAGAGGACACAGGCGCAGAGATGCTTTTGTGCT 2411
Dh 2755 TCCGTGACACACAGCGCAGATCAGAGGACACAGGCGCAGAGATGCTTTTGTGCT 2814
QY 2412 TTCTGGCCTGTGGTACGATTTTTCATGAGCTTACAGATGCTGCTGCTGCTGCTGCTGCT 2471
Dh 2815 TTCTGGCCTGTGGTACGATTTTTCATGAGCTTACAGATGCTGCTGCTGCTGCTGCTGCT 2534
QY 2472 ATTAACACATTTTCATATTTGTTA 2497
Dh 2875 ATTAACACATTTTCATATTTGTTA 2900
RESULT 4
BC004176 1905 bp mRNA linear PRI 12-JUL-2001
LOCUS Homo sapiens, similar to hypothetical protein FLJ10928, clone
DEFINITION MG:2772 IMAGE:2958967, mRNA, complete cds.
ACCESSION BC004176
VERSION BC004176.1 GI:13278815

QY 1994 GTCGCTGCTCAAGGCGCTCAGACCTTTCTAACGCGATGTGTAGAGGAGCTGACGTA 2053
|||||
Db 1381 GTCGCTGCTCAAGGCGCTCAGACCTTTCTAACGCGATGTGTAGAGGAGCTGACGTA 1440
QY 2054 CTTTGGGGGCAACAGCAGCCTAGTTTCTATCTCAACTAGCCCTGACACTACCTGT 2113
|||||
Db 1441 CTTTGGGGGCAACAGCAGCCTAGTTTCTATCTCAACTAGCCCTGACACTACCTGT 1500
QY 2114 GGACAGGAATGAAACAGACCTTCCGTCGAAAGAGGATACGCTCCACCCCGCC 2173
|||||
Db 1501 GGACAGGAATGAAACAGACCTTCCGTCGAAAGAGGATACGCTCCACCCCGCC 1560
QY 2174 CTTCTGACACTCTCTCTCTCTCCAGTTCTCTGGAACAGCAGGAGGAGCAACCA 2233
|||||
Db 1561 CTTCTGACACTCTCTCTCTCTCCAGTTCTCTGGAACAGCAGGAGGAGCAACCA 1620
QY 2234 GTGGCCCCCAAGAGGAGGAGATCTCTAGGCCCCAGCCGGGAGGCTGGAAGGCGTGG 2293
|||||
Db 1621 GTGGCCCCCAAGAGGAGGAGATCTCTAGGCCCCAGCCGGGAGGCTGGAAGGCGTGG 1680
QY 2294 CAGATGCTTCCCTCATCCACCTCCACCGGTCAGGCTTTGCTGCTCCGACACCTC 2353
|||||
Db 1681 CAGATGCTTCCCTCATCCACCTCCACCGGTCAGGCTTTGCTGCTCCGACACCTC 1740
QY 2354 CTTGTACACACGAGCAGATCAGAGGCGCAGGAGGAGATAGTCTTTTGTCTTT 2413
|||||
Db 1741 CTTGTACACACGAGCAGATCAGAGGCGCAGGAGGAGATAGTCTTTTGTCTTT 1800
QY 2414 CTTGGCTCTGCTAGTACATTTTCTAGCTTCTAGCTTCTGCTTTTGTCTTACTGAGAAAT 2473
|||||
Db 1801 CTTGGCTCTGCTAGTACATTTTCTAGCTTCTAGCTTCTGCTTTTGTCTTACTGAGAAAT 1860
QY 2474 AAAACACATTTTCATATTTGTTAAAAA 2518
|||||
Db 1861 AAAACACATTTTCATATTTGTTAAAAA 1905
RESULT 5
BC004210 1905 bp mRNA linear PRI 12-JUL-2001
LOCUS
DEFINITION Homo sapiens, similar to hypothetical protein FLJ10928, clone
ACCESSION BC004210
VERSION BC004210.1 GI:13278902
KEYWORDS MGC.
SOURCE
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (01-MAR-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK
COMMENT NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: gcaps-remail.nih.gov
Tissue Procurement: DCTD/DRP
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
http://www.systemsbio.org
contact: amadan@systemsbiology.org
Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia
Greene, Mark Kettelman and Anuradha Madan
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
Series: IRAL Plate: 10 Row: 9 Column: 1.
Location/Qualifiers
1. 1905

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:4436 IMAGE:2958967"
/tissue_type="Kidney, renal cell adenocarcinoma"
/clone_lib="NIH_MGC_14"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
272. 1171
/codon_start=1
/product="Similar to hypothetical protein FLJ10928"
/protein_id="AAH04210.1"
/db_xref="GI:13278903"
/translation="MAREIDNVEPERFTYHNVRLWDEESADLPKMEHREHTEARA
OETHLVHCKMGVSRSAATVLAAMKQYECSEALRHOELRPIARPPEFLQLOI
YOGILJASRSHWEOKYGVSPHEHPAPESTPPPLPPEEGGEEKVYGMESOA
APKEEPGRPRINLEKGVRSISLEPSELESTSETPSPPEVSSHESEHPLOP
OLATFGGOVDROGPALPKRSQSVTLQGSAYVANRTOARQEDQSGGGEPCIS
SIPRRKVVYRQASVHDSGEERA"
BASE COUNT 418 a 627 c 517 g 343 t
ORIGIN
Query Match 64.6%; Score 1641; DB 9; Length 1905;
Best Local Similarity 94.3%; Pred. No. 0;
Matches 1797; Conservative 0; Mismatches 5; Indels 103; Gaps 5;
QY 712 GAGCAGGCGATCCGTCGCTAGCTGTGGAAGATGTTGATGTCAGTACCTGAGAGTGC 771
|||||
Db 6 GAGGAGGCGATCCGTCGCTAGCTGTGGAAGATGTTGATGTCAGTACCTGAGAGTGC 64
QY 772 ACTTCCAAAGAGATCGGCGAGGCTGTGGAGTGGGCGCTGGGGCTCCCTCCAGAGTAC 831
|||||
Db 65 ACTTCCAAAGAGATCGGCGAGGCTGTGGAGTGGGCGCTGGGGCTCCCTCCAGAGTAC 124
QY 832 CGTGACTTCATCGACACAGATGCTGCTGTGTGGACAGACGCGGAGCGAGCTCCCGC 891
|||||
Db 125 CGTGACTTCATCGACACAGATGCTGCTGTGTGGACAGACGCGGAGCGAGCTCCCGC 184
QY 892 ATCTTCCCGCACCTCTACCTGGGCTCAGAGTGAAGAGGAGGAGAACTGTGAGAGTGCAG 951
|||||
Db 185 ATCTTCCCGCACCTCTACCTGGGCTCAGAGTGAAGAGGAGGAGAACTGTGAGAGTGCAG 244
QY 952 AGGAACAGGCTACCCACATCTTGAACATGGCCCGGAGATTGAACATCTTACCTTGA 1011
|||||
Db 245 AGGAACAGGCTACCCACATCTTGAACATGGCCCGGAGATTGAACATCTTACCTTGA 304
QY 1012 CGCTTACCTTACCAATGTGCGCTTGGGATGAGAGTGGCCCACTCTGCGCAC 1071
|||||
Db 305 CGCTTACCTTACCAATGTGCGCTTGGGATGAGAGTGGCCCACTCTGCGCAC 364
QY 1072 TGAAGAGAGCGACCGCTTATTTAGAGCTCAAGAGACAGGCGACCCACCTGTGGTC 1131
|||||
Db 365 TGAAGAGAGCGACCGCTTATTTAGAGCTCAAGAGACAGGCGACCCACCTGTGGTC 424
QY 1132 CACTGCAAGATGGGCGCTACGCGCTCAGCGGCGACAGTGTGGCTTATGCCATGAACAG 1191
|||||
Db 425 CACTGCAAGATGGGCGCTACGCGCTCAGCGGCGACAGTGTGGCTTATGCCATGAACAG 484
QY 1192 TACGAATGACGCTGGAGGAGGCGCTGGCGACGTGAGAGAGTCTCCGGCCATGGCCGC 1251
|||||
Db 485 TACGAATGACGCTGGAGGAGGCGCTGGCGACGTGAGAGAGTCTCCGGCCATGGCCGC 544
QY 1252 CCGAACCCTGCTTCTGCGGCGACGCTCAGATCTACAGGAGCATCTGACG----- 1302
|||||
Db 545 CCGAACCCTGCTTCTGCGGCGACGCTCAGATCTACAGGAGCATCTGACGCGGCGCGC 604
QY 1303 ----- 1302
Db 605 CAGAGCATGTCTGGAGAGAAAGTGGGTCTCCCAAGAGAGCACCCACCCCT 664
QY 1303 -----GCCAGACCTGAGGCTGTGGGAGGAGAG 1333
Db 665 GAGTCTCTACACCATTTCCACCTCTTCCGCGACAGACTGAGGCTGTGGGAGGAGAG 724


```

Db 541 TTCAGCGTGCAGCTGTGTGGGCAAGCCGATCTTCAAGCCCATCTCCATCCAGACCATG 600
Qy 688 T-----
Db 601 TGGGCCACACTCCAGGTATTGACCAAGCATGTGAGGACACTCTAGGCGCGGCTTGA 660
Qy 689 -----
Db 661 CGGGGTGCAGTGTCCCTCACCTGGGCGACCCACTACAGGAGAGACTGAACTCCAGACAG 720
Qy 689 -----
Db 721 AACTGCTCAATGATGAGCGGCTATGGCCGACTGTGAGTCTCTGGGCTCCAGCGCC 780
Qy 689 -----GGTCCCTCAGAAACAGAGAGAGAGATGAGAGAGGCGATCCGTGTGAGTGTGG 738
Db 781 GAGCTGTGGGGGTCTTCAGAAACAGAGAGAGATGAGAGAGGATCCGTGTGAGTGTGG 840
Qy 739 AAGGTGTGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 798
Db 841 AAGGTGTGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
Qy 799 GAGTGTGCTGTGGGCTTCCCTCCAGCATGATCCGCTTCCCTCCAGCATGATCCGCTTCC 858
Db 901 GAGTGTGCTGTGGGCTTCCCTCCAGCATGATCCGCTTCCCTCCAGCATGATCCGCTTCC 960
Qy 859 CTGCTGTGTGAGCAGAGCGGAGCGAGCTCCGCTTCCCTCCAGCATGATCCGCTTCCAGTCA 918
Db 961 CTGCTGTGTGAGCAGAGCGGAGCGAGCTCCGCTTCCCTCCAGCATGATCCGCTTCCAGTCA 1020
Qy 919 GAGTGTGAGCAGAGCAGCAGTGTGAGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 978
Db 1021 GAGTGTGAGCAGAGCAGCAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
Qy 979 ATGGCCCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1038
Db 1081 ATGGCCCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
Qy 1039 TGGGATGAGAGTGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1098
Db 1141 TGGGATGAGAGTGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
Qy 1099 GCTGTGAGAGACACAGAGGACCCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1158
Db 1201 GCTGTGAGAGACACAGAGGACCCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1260
Qy 1159 GCGGCCACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1218
Db 1261 GCGGCCACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1320
Qy 1219 GCGCACGTGACAGAGCTCGGCGCATCGGCGCATCGGCGCATCGGCGCATCGGCGCATCGG 1278
Db 1321 GCGCACGTGACAGAGCTCGGCGCATCGGCGCATCGGCGCATCGGCGCATCGGCGCATCGG 1380
Qy 1279 CAGATCTTACAGAGGATCTCTGACG----- 1302
Db 1381 CAGATCTTACAGAGGATCTCTGACGCGCCAGAGCCATGTCTGGGAGCAGAAAGTG 1440
Qy 1303 ----- 1302
Db 1441 GGTGGGTCTCTCCAGAGAGAGACCCCTGAAGTCTTACACCATTTCCACTCTT 1500
Qy 1303 --GCCAGAACTGAGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1360
Db 1501 CCGCCAGAACTGAGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1560
Qy 1361 GCGCCGAAAGAGAGCTGTGGGCGCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1420
Db 1561 GCGCCGAAAGAGAGCTGTGGGCGCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1618
Qy 1421 CCATCAGTCTTGTGAGAGCTCTTGTGAGCTGTGAGAGACCTTCAATGAGAGCAGTGTGAT 1480
Db 1619 CCATCAGTCTTGTGAGAGCTCTTGTGAGCTGTGAGAGACCTTCAATGAGAGCAGTGTGAT 1676

```

```

Qy 1481 GCGAGAGCTCTTCTCTTCCACAGAGTCTTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAG 1540
Db 1677 GCGAGAGTCTTCTCTTCTTCCACAGAGTCTTCAATGAGAGAGAGAGAGAGAGAGAGAGAG 1736
Qy 1541 GCTTGCAGAGCAGCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1600
Db 1737 GCTTGCAGAGCAGCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1796
Qy 1601 CCGCCAGTCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1660
Db 1797 CCGCCAGTCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1856
Qy 1661 CAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1720
Db 1857 CAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1916
Qy 1721 CAGGTTCCGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1780
Db 1917 CAGGTTCCGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1976
Qy 1781 CTGA 1784
Db 1977 CTGA 1980

RESULT 7
AX398768 1755 bp DNA linear PAT 27-MAY-2002
LOCUS AX398768
DEFINITION Sequence 4 from Patent WO0220732.
ACCESSION AX398768
VERSION AX398768.1 GI:21261303
KEYWORDS
SOURCE
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1
AUTHORS Liou,J.R.
TITLE Regulation of human map kinase phosphatase-like enzyme
JOURNAL Patent: WO 0220732-A 4 14-MAR-2002;
Bayer Aktiengesellschaft (DE)
FEATURES
source
1..1755
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 350 a 577 c 505 g 323 t
ORIGIN

Query Match 51.3%; Score 1303.6; DB 6; Length 1755;
Best Local Similarity 94.2%; Pred. No.2,4e-258;
Matches 1430; Conservative 0; Mismatches 4; Indels 84; Gaps 4;

Qy 979 ATGGCCCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1038
Db 322 ATGGCCCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 381
Qy 1039 TGGGATGAGAGTGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1098
Db 382 TGGGATGAGAGTGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 441
Qy 1099 GCTGTGAGAGCAGAGGCGACACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1158
Db 442 GCTGTGAGAGCAGAGGCGACACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 501
Qy 1159 GCGGCCACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1218
Db 502 GCGGCCACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 561
Qy 1219 GCGCACGTGACAGAGCTCGGCGCATCGGCGCATCGGCGCATCGGCGCATCGGCGCATCGG 1278
Db 562 GCGCACGTGACAGAGCTCGGCGCATCGGCGCATCGGCGCATCGGCGCATCGGCGCATCGG 621

```

```

Qy 1279 CAGATCTACAGAGGATCTCTGACGGCCAGAACCTGAGGTGGTGGAGAGAGAGTTGT 1338
    |||
Db 622 CAGATCTACAGAGGATCTCTGACGGCCAGAACCTGAGGGTGGTGGAGAGAGAGTTGT 681
Qy 1339 AGGATGGAAGAGACCCAGAGCCCGAAGAGAGAGCTGGGGCCAGGGGACAGTATA 1398
    |||
Db 682 AGGATGGAAGAGACCCAGAGCCCGAAGAGAGAGCTGGGGCCAG - GGCCAGTATA 739
Qy 1399 AACCCCGAGGGGTATGATGGGTCCATGATCTCTGAGAGCCCTCTGGAGCTGGAGAG 1458
    |||
Db 740 AACCTCCAGGGGTATGATGGGTCCATGATCTCTGAGAGCCCTCT - GGAGCTGGAGAG 798
Qy 1459 CACCTCAAGTACAGATGACATGACAGAGTCTTCTTCCACAGATCTTCAATGAAG 1518
    |||
Db 799 CACCTCAAGTACAGATGACATGACAGAGTCTTCTTCCACAGATCTTCAATGAAG 857
Qy 1519 AGCTCTGACAGCCCTTCCACAGCTTTCAGAGACCAAGGAGGACAGAGTGGACAGGG 1578
    |||
Db 858 AGCTCTGACAGCCCTTCCACAGCTTTCAGAGACCAAGGAGGACAGAGTGGACAGGG 917
Qy 1579 GGCTCAGAGCTGCGGTGAAGTCCCGCAGTCAAGTGTACCTCCAGAGGACAGTGGCGTGG 1638
    |||
Db 918 GGCTCAGAGCTGCGGTGAAGTCCCGCAGTCAAGTGTACCTCCAGAGGACAGTGGCGTGG 977
Qy 1639 TGGCCACCGGACCCAGAGGCTTTCAGAGACAGAGGAGGAGGAGGAGGAGGAGGAG 1698
    |||
Db 978 TGGCCACCGGACCCAGAGGCTTTCAGAGACAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1037
Qy 1699 AGCTCTGATTTCTCTACAGCCAGGTTCCGAGAGTGTGAGACAGGACAGGCTGTCATG 1758
    |||
Db 1038 AGCTCTGATTTCTCTACAGCCAGGTTCCGAGAGTGTGAGACAGGACAGGCTGTCATG 1097
Qy 1759 ACAGTGAAGAGAGGAGGAGGAGGCTGAGGCTCAGACATGCCAGCTCCCTGACACATGA 1818
    |||
Db 1098 ACAGTGAAGAGAGGAGGAGGAGGCTGAGGCTCAGACATGCCAGCTCCCTGACACATGA 1157
Qy 1819 AGAGATCCACAATCTCTTGAGAAACACCTCAGCTGTGTGCCGACACATCTCTCTTC 1878
    |||
Db 1158 AGAGATCCACAATCTCTTGAGAAACACCTCAGCTGTGTGCCGACACATCTCTCTTC 1217
Qy 1879 AGCTCGGCGCCATACCTGCTACATACAGCTTACCTCCACCCCTGCTACATAGGCTCAG 1938
    |||
Db 1218 AGCTCGGCGCCATACCTGCTACATACAGCTTACCTCCACCCCTGCTACATAGGCTCAG 1277
Qy 1939 CTCCACCCCTGCTACATACAGCTTACCTCCTTACAGCTTAAAGTCCAGAGGACATGTG 1998
    |||
Db 1278 CTCCACCCCTGCTACATACAGCTTACCTCCTTACAGCTTAAAGTCCAGAGGACATGTG 1337
Qy 1999 CCTGTCAAGGGCTCAGACATTTCTACTGGATGTGTAGAGGAGTGAAGTACTTT 2058
    |||
Db 1338 CCTGTCAAGGGCTCAGACATTTCTACTGGATGTGTAGAGGAGTGAAGTACTTT 1397
Qy 2059 GGGGGCAACAGACCTTACTTCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTT 2118
    |||
Db 1398 GGGGGCAACAGACCTTACTTCTTCACTTCTCACTTCTCACTTCTCACTTCTCACTTCT 1434
Qy 2119 GGAATGAAAAACAGAGCTTCCCGTGGAAAAAGGTCACGCGCTCCACCCCGCCCTCC 2178
    |||
Db 1435 -----CCC 1437
Qy 2179 TGCACCTCTCTCTCTCTCCAGATTCTCTGGAACACAGCCAGGCGAGCAACAGTGGC 2238
    |||
Db 1438 TGCACCTCTCTCTCTCTCCAGATTCTCTGGAACACAGCCAGGCGAGCAACAGTGGC 1497
Qy 2239 CCCCAAAAGGAGGAGATCTCTCAGAGCCCGGAGGAGCTGGAAGGCTGGCAGAT 2298
    |||
Db 1498 CCCCAAAAGGAGGAGATCTCTCAGAGCCCGGAGGAGCTGGAAGGCTGGCAGAT 1557
Qy 2299 CGCTTCCCTCATCTCAGCTCAGCGGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 2358
    |||
Db 1558 CGCTTCCCTCATCTCAGCTCAGCGGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 1617
Qy 2359 ACACCAAGCAGATCAGAGGAGCAGAGGCTCAGATAGTCTTTTGTCTTCTTGTGCTTGTGCT 2418

```

```

Db 1618 ACACCAAGCAGATCAGAGGAGCAGAGGAGATAGTCTTCTTGTCTCTTCTGCGC 1677
Qy 2419 CTCTGGCTAGTACAGTTTTCATAGCTTACATCTGCTTGTACTGAGAAATAAAC 2478
    |||
Db 1678 CTCTGGCTAGTACAGTTTTCATAGCTTACATCTGCTTGTACTGAGAAATAAAC 1737
Qy 2479 ACATTTTCATTTGGTT 2496
    |||
Db 1738 ACATTTTCATTTGGTT 1755

RESULT 8
AK001790
LOCUS
DEFINITION
    AK001790 1755 bp mRNA linear PRI 01-AUG-2002
Homo sapiens CDNA FLJ10928 fls, clone OVARC1000473, weakly similar
to DUAL SPECIFICITY PROTEIN PHOSPHATASE 3 (EC 3.1.3.48) (EC
3.1.3.16).
ACCESSION
    AK001790
VERSION
    AK001790.1 GI:7023282
KEYWORDS
    oligo capping, fls (full insert sequence).
SOURCE
    Homo sapiens ovarian cancer CDNA to mRNA, clone_11b:OVARC1
    clone:OVARC1000473.
ORGANISM
    Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
    1
    Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
    Nishikawa,T., Nagai,K., Sugano,S., Ishibashi,T., Fujimori,K.,
    Tanai,H., Kimata,M., Watanabe,M., Hirooka,S., Ishii,S., Kawai,Y.,
    Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagahara,K.,
    Masuho,Y., and Kanehori,K.
    NEDO human cDNA sequencing project
    Unpublished
    2 (bases 1 to 1755)
    Isogai,T. and Otsuki,T.
    Direct Submission
    Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute,
    Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
    (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
    NEDO human cDNA sequencing project supported by Ministry of
    International Trade and Industry of Japan; CDNA full insert
    sequencing: Research Association for Biotechnology; CDNA library
    construction, 5'- & 3'-end one pass sequencing and clone selection:
    Helix Research Institute (supported by Japan Key Technology Center
    etc.) and Department of Virology, Institute of Medical Science,
    University of Tokyo.
FEATURES
    source
        Location/Qualifiers
            1..1755
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="OVARC1000473"
                /tissue_type="ovarian cancer"
                /clone_11b="OVARC1"
                /note="Cloning vector: PME18SFL3"
                232..657
                /note="unnamed protein product"
                /codon_start=1
                /protein_id="BA091913.1"
                /db_xref="GI:7023283"
                /translation="MEGTMAMQORPVLSSQHPFVLNSSPAHSPAREINDFYPERFT
                YHNYLMEESAOQLPHWKETHHRELEAARAGTIVLVCKMGVSRSATVLAVMKQY
                ECSLEQALRHQELRPTRAPRPGFLROLOIYQILTART"
BASE COUNT
    350 a 577 c 505 g 323 t
ORIGIN
Query Match 51.3%; Score 1303.6; DB 9; Length 1755;
Best Local Similarity 94.2%; Pred. No. 2.4e-258;
Matches 1430; Conservative 0; Mismatches 4; Indels 84; Gaps 4;
Qy 979 ATGGCCCGGAGATTCAGACACTTCTACCTGAGCGCTTACCTACACACATGTGGCCTC 1038
    |||
Db 322 ATGGCCCGGAGATTCAGACACTTCTACCTGAGCGCTTACCTACACACATGTGGCCTC 381

```


Db 383 GGAGAGAGAGGAGCACTGCAAACTCATAGAGCGCTGAGGCCGAGGATGACAT 442
 Oy 420 CCGGCTGGCA--GGCCACTGGAGGACACCGGCGCTCCCGGCTCCGTACTGCG----- 472
 Db 443 CCGACTTGGAGACCACTGGAGGACACCGGCGCTCCCGGCTCCCGGCTCCGTACTGCT 502
 Oy 473 TGTAGTTTCTACAGAGAAAG-GAGAAAGTCTGAG-CCAGATGAGAGGCTCTCTGG 530
 Db 503 TGTAGTTTCTACAGAGAAAGGAGTCTGAGGCCAGATGAGAGGCTCTCTCTGG 562
 Oy 531 -CGTGGATTTCTGACAGCAGCTCCCGAGCTGACACCTGGGCTGGTCTTGGCCCTCT 589
 Db 563 ACGTGGATTTCTGACAGCAGCTCCCGAGCTGACACCTGGGCTGGTCTTGGCCCTCT 622
 Oy 590 GGAGTGAACCCAGGTGACTTAGATGAGAGCGGGGCTTACAGCTGACGCTGGTGGC 649
 Db 623 GGAGTGAACCCAGGTGACTTAGATGAGAGCGGGGCTTACAGCTGACGCTGGTGGC 682
 Oy 650 AAAGCCGGATCTTCAAGCCCATCTCCATCCAGACCATGT----- 688
 Db 683 AAAGCCGGATCTTCAAGCCCATCTCCATCCAGACCATGTGGGCCACACTCCAGGATTTGC 742
 Oy 689 ----- 688
 Db 743 ACCAAGCATGTGAGGACGCTTAGGACAGCGGCTTGTACGGGTGGAGTGCCTCACCT 802
 Oy 689 ----- 688
 Db 803 GGGCCAGCACTACAGAGAGAGACTGAACCTCGAAGAGCTGCCCTCAATGAGTGAAG 862
 Oy 689 -----GGTCTCAGAAC 700
 Db 863 CTATGGCCGAGCTGAGAGTCTTGGGGCTCCAGAGCGCCAGACCTGGCGGTCTTAGAAC 922
 Oy 701 AGGAGCAGATGAGAGAGAGCGATCCGTGCTGAGCTGTGAAAGTGTGGATGTCAGTGAC 760
 Db 923 AGGAGCAATGAGAGAGAGCGATCCGTGCTGAGCTGTGAAAGTGTGGATGTCAGTGAC 982
 Oy 761 TGGAGAGTGTCACTTCCAAAGAGATCCCGCAGGCTGTGAGCTGGCGCTGGGGCTCCCC 820
 Db 983 TGGAGAGTGTGCTTCCAGAGAGATCCCGCAGGCTGTGAGCTGGCGCTGGGGCTCCCC 1042
 Oy 821 TCCAGCAGTACCTGATCTCATTCAGAACACAGATGCTGCTGTGTGACAGCGGGAC 880
 Db 1043 TCCAGCAGTACCTGATCTCATTCAGAACACAGATGCTGCTGTGTGACAGCGGGAC 1102
 Oy 881 GAGGCTCCGCACTTCCCAACCTCTACCTGGGCTAGAGTGAACGACCAAACTGG 940
 Db 1103 GAGGCTCCGCACTTCCCAACCTCTACCTGGGCTAGAGTGAACGACCAAACTGG 1162
 Oy 941 AGGAGCTGAGAGAGAGAGGCTCAACCCATCTTGAACATGCGCCGGAGATGCAACT 1000
 Db 1163 AGGAGCTGAGAGAGAGAGGCTCAACCCATCTTGAACATGCGCCGGAGATGCAACT 1219
 Oy 1001 TCTACCTGAGAGGCTTCACTACCAACAATGCGGCTGTGGATGAGAGCGGGAC 1060
 Db 1220 TCTACCTGAGAGGCTTCACTACCAACAATGCGGCTGTGGATGAGAGCGGGAC 1279
 Oy 1061 TGTGCGCGCACTGGAAGAGAGACCGCTTCAATGAGCTGCAAGAGACAGGCGACCC 1120
 Db 1280 TGTGCGCGCACTGGAAGAGAGACCGCTTCAATGAGCTGCAAGAGACAGGCGACCC 1339
 Oy 1121 ACGTGTGCTGCTACGAAAGATGGGCTGACAGCGCTGACGCGCCACAGTGTGGCTATG 1180
 Db 1340 ACGTGTGCTGCTACGAAAGATGGGCTGACAGCGCTGACGCGCCACAGTGTGGCTATG 1399
 Oy 1181 CCATGAAGCACTAGCAATGACGCTGAGAGCGGCTGCGCAGCTGAGAGAGCTCCGGC 1240
 Db 1400 CCATGAAGCACTAGCAATGACGCTGAGAGCGGCTGCGCAGCTGAGAGAGCTCCGGC 1459
 Oy 1241 CCATGCGCGCGCCCAACCTGGCTTCTGCGCAGCTGACAGTCTACAGGGAATCTCTGA 1300
 |||||||

Db 1460 CCATGCGCGCGCCCAACCTGGCTTCTGCGCAGCTGACAGATCTACAGGGAATCTGA 1519
 Oy 1301 CG----- 1302
 Db 1520 CGGCCAGCGCCAGAGCCATGTCTGGGAGACAGAAAGTGGGTGGTCCCGAGAGAGC 1579
 Oy 1303 -----GCCAAGACCTGAGGGTGGT 1322
 Db 1580 ACCGAGCCCTGAAAGTCTACACCATTTCCACCTCTCCGCGAAGACCTGAGGGTGGT 1639
 Oy 1323 GGGAGAGAAAGTTGTAGCATGAGAGAGAGCCAGAGCCCGCAAGAGAGCTGGGG 1382
 Db 1640 GGGAGAGAAAGTTGTAGCATGAGAGAGAGCCAGAGCCCGCAAGAGAGCTGGG- 1698
 Oy 1383 CCAGGGGACAGTAAACCTCCGAGGGGTGATGAGTGCATGATCTTGGAGACCTTC 1442
 Db 1699 -CCAGGGGACAGTAAACCTCCGAGGGGTGATGAGTGCATGATCTTGGAGACCTTC 1757
 Oy 1443 CTTGGAGAGCTGAGAGACCTCAGTAGACACAGTGCATGAGAGAGTCTTCTTCCAC 1502
 Db 1758 CTT-GGAGCTGAGAGAGACCTCAG-AGACAGTGCATGAGAGAGTCTTCTTCCAC 1815
 Oy 1503 GAGTCTTCAATGAAGAGAGCTCTGAGAGCTTCCACAGCTTGAAGAGCAAGGAGGC 1562
 Db 1816 GAGTCTTCAATGAAGAGAGCTCTGAGAGCTTCCACAGCTTGAAGAGCAAGGAGGC 1875
 Oy 1563 CAGCAGGTGAGAGAGGGGCTCAGCTGCTGAAAGTCCCGCAGTCAAGTGTTCCTC 1622
 Db 1876 CAGCAGGTGAGAGAGGGGCTCAGCTGCTGAAAGTCCCGCAGTCAAGTGTTCCTC 1935
 Oy 1623 CAGGCACTGCGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1682
 Db 1936 CAGGCACTGCGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1995
 Oy 1683 GGGCAGGGGCAAGGAG 1742
 Db 1996 GGGCAGGGGCAAGGAG 2055
 Oy 1743 CAGGCACTGAG 1802
 Db 2056 CAGGCACTGAG 2113
 Oy 1803 GCTCCCTGAG 1862
 Db 2114 GCTCCCTGAG 2173
 Oy 1863 CGCACATTTCTCTCAGCTCGGCCCATTTACCGCTGCTACAGCTCACTCCACCTT 1922
 Db 2174 CGCACATTTCTCTCAGCTCGGCCCATTTACCGCTGCTACAGCTCACTCCACCTT 2233
 Oy 1923 GTCACTAGGGGCTGACCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1982
 Db 2234 GTCACTAGGGGCTGACCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2293
 Oy 1983 CCAAGGCCCATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2011
 Db 2294 CCAAGGCCCATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2322

RESULT 10
 AK000522
 LOCUS Homo sapiens cDNA FLJ20515 fis, clone KAT09889.
 DEFINITION Homo sapiens cDNA FLJ20515 fis, clone KAT09889.
 ACCESSION AK000522
 VERSION AK000522.1 GI:7020673
 KEYWORDS oligo capping; fis (full insert sequence).
 SOURCE Homo sapiens signet-ring cell carcinoma cell_line:KATO III cDNA to mRNA, clone, lib:KAT clone:KAT09889.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (sites)
 AUTHORS Matanabe,K., Kumagai,A., Itakura,S., Yamazaki,M., Tashiro,H.,

Ota, T., Suzuki, Y., Obayashi, M., Nishi, T., Shibahara, T., Tanaka, T.,
 Nakamura, Y., Isogai, T. and Sugano, S.
 NEDO human cDNA sequencing project
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 Submitted (15-FEB-2000) Sumio Sugano, Institute of Medical Science,
 University of Tokyo, Department of Virology, Shirokane-dai, 4-6-1,
 Minato-ku, Tokyo 108-8639, Japan (E-mail: cdna@ims.u-tokyo.ac.jp,
 tel: 81-3-5449-5286, Fax: 81-3-5449-5416)
 NEDO human cDNA sequencing project supported by Ministry of
 International Trade and Industry of Japan: cDNA full insert
 sequencing: Research Association for Biotechnology: cDNA library
 construction, 5'- & 3'-end one pass sequencing: Department of
 Virology and Human Genome Center, Institute of Medical Science,
 University of Tokyo (partly supported by Science and Technology
 Agency).

FEATURES
 source
 1. 2064
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="KAT09889"
 /cell_line="KATO III"
 /cell_type="signet-ring cell carcinoma"
 /clone_id="Kat"
 /note="cloning vector pME18SFL3"
 147. 1331
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="BAA9128.1"
 /db_xref="GI:7020674"
 /translation="MALVYVSRSPGSGASRPVGPMDAVORRSRLORRSPNAVLRGA
 VGLDDGGDNDAAEASPEPEKAPSEELHGDHDFDGGSPKQEPORHLHLMV
 QLLRPDDIRLAOLEAPRPRLRYLIYVREGSLSDDELVLGVPPDSSTCT
 LGVLVMSDPQVYLDGCGSVYSGGSRIRKPISTGMATGLVLAQCEAAAGSG
 LIPGSAITWASHYDERLNSDSCLENTWMAADLESLPPSAPEGSSSEORMEQAIR
 AELMKVLELESTSETSDMEYFSSHESHEPLPFPQLARTKGQOYDRGROPALRS
 ROSVVTLDGSAVAVNRTOAFQEOEGOGGCGEPCISSTPRRKVVRASVHDSSEBS
 EA"

CDS
 BASE COUNT 421 a 671 c 605 g 367 t
 ORIGIN

Query Match 48.2%; Score 1224.2; DB 9; Length 2064;
 Best Local Similarity 74.7%; Pred. No. 5.7e-242;
 Matches 1881; Conservative 0; Mismatches 123; Indels 513; Gaps 7;

Db 420 GAGGAGCAGAGGCGACACCTGACCTCATGTGTACAGTCTGTGAGCCGCGAGATGACATC 479
 QY 421 CGCTGGCAGCCCGACGCTGGAGGCAACCCGGCTCTCCGGCTCGCTACCTGCTGTAGTTT 480
 Db 480 CGCTGGCAGCCCGACGCTGGAGGCAACCCGGCTCTCCGGCTCGCTACCTGCTGTAGTTT 539
 QY 481 TCTACAG 540
 Db 540 TCTACAG 599
 QY 541 CTGACAGCAGCTCCCGACCTGACCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
 Db 600 CTGACAGCAGCTCCCGACCTGACCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 659
 QY 601 CAGGTGACTTATGATGAG 660
 Db 660 CAGGTGACTTATGATGAG 719
 QY 661 TTGAGCCCATCTCATCCAGACCATGTGCTCTCGAAGAGAGAGAGAGAGAGAGAGAGAGAG 720
 Db 720 TTGAGCCCATCTCATCCAGACCATGTGCTCTCGAAGAGAGAGAGAGAGAGAGAGAGAGAG 762
 QY 721 ATCCGCTGAGCTGAGAAAGTGTGATGTGATGTGATGTGATGTGATGTGATGTGATGTGAT 763
 Db 763 -----
 QY 781 GAGATCCGACAGGCTGTGAGAGCTGCGCTGGGGCTCCCGCTCCAGAGATACCGTACCTTC 840
 Db 781 -----
 QY 763 -----
 Db 763 -----
 QY 841 ATCGACACACCATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
 Db 841 -----
 QY 769 ACCAACAATGTAGGACAGCTCTAGGACAGGCGCTGTACCGG 811
 Db 901 CACCTCTACCTGGGCTCAGAGTGAAGCAGAGAACTCGAGAGAGCTCAGAGAGAGAGAGAG 960
 Db 812 -----
 QY 961 GTACCCACATCTTGAACATGAGCCCGGAGATTGACAACTTACCTGAGCGCTTACC 1020
 Db 812 -----
 QY 1021 TACCAATATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
 Db 858 TCCGACAGAGCTGCTCAATGA 880
 QY 1081 ACGACCGCTTCAATTGAGGCTGCAAGACACAGGACACCAAGTGTGCTGCTGCTGCTGCTGCTG 1140
 Db 881 -----
 QY 1141 ATGGCGCTCAGCCGCTCAGCGGCCACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
 Db 881 -----
 QY 1201 AGCTGAGACAGCCCTGCGCCACAGTGCAGAGAGCTCCGCGCCATGCGCGGCCACCACTT 1260
 Db 881 -----
 QY 1261 GCGTTCCTGCGCCCGACGCTGCAAGATCTACAGGCGCATCTGACGCGCCAGAACCTGAGG 1320
 Db 936 GGGTCTCTAGAACAGGAGCAGATGTGACAGCGCATCC -----GTGC 976
 QY 1321 TGGGAG 1380
 Db 977 TGAGCTGTGGAAT 991
 QY 1381 GCGCAGCGGCGACGTATAAAGCTCCAGAGGATGATGATGATGATGATGATGATGATGATGAT 1440
 Db 992 -----
 QY 1441 TCCCTGGAGCTGAG 1500

```
Dh 992 ---GTTGAGCTGGAGACACCTCTAG-AGACAGTGAATGCCAGAGGCTCTCTCTCC 1047
Oy 1501 AGAGTCTTCACATGAAGAGCCTCGACGCCCTTCCACAGCTTGCAGAGAACAGGAG 1560
Dh 1048 ACAGTCTTCACATGAAGAGCCTCTGACGCCCTTCCACAGCTTGCAGAGAACAGGAG 1107
Oy 1561 GCCAGCAGGTGGACAGAGGGGCTCAGCCTGCCCTGAAAGTCCCGCAGTCAAGTGTACCC 1620
Dh 1108 GCCAGCAGGTGGACAGAGGGGCTCAGCCTGCCCTGAAAGTCCCGCAGTCAAGTGTACCC 1167
Oy 1621 TCCAGAGCAGTGGCTGTGTGGCCCAACCGGACCCAGCCTTCCAGAGAGAGAGAGAGGAG 1680
Dh 1168 TCCAGAGCAGTGGCTGTGTGGCCCAACCGGACCCAGCCTTCCAGAGAGAGAGAGAGGAG 1227
Oy 1681 AGGAGCAGAGGAGAGAGAGCCTCATTTCTCTACGCCAGTCCGAGTCCGAGAGAGTGA 1740
Dh 1228 AGGAGCAGAGGAGAGAGAGCCTCATTTCTCTACGCCAGTCCGAGTCCGAGAGTGA 1287
Oy 1741 GACAGGCCAGGCTGATGATGAGAGAGAGAGAGGAGGAGGCTGAGCCCTGACACATGCC 1800
Dh 1288 GACAGGCCAGGCTGATGATGAGAGAGAGAGAGGAGGAGGCTGAGCCCTGACACATGCC 1347
Oy 1801 ACCTCTCTCTACACTGAAGAGATCCACAACTCTTGGAGAAACACCTTCACTCTGT 1860
Dh 1348 ACCTCTCTCTACACTGAAGAGATCCACAACTCTTGGAGAAACACCTTCACTCTGT 1407
Oy 1861 GCCGACACATTTCTCTAGCTCCGCCCATATCCGCTCACTACAGCTCACTCCACCC 1920
Dh 1408 GCCGACACATTTCTCTAGCTCCGCCCATATCCGCTCACTACAGCTCACTCCACCC 1467
Oy 1921 CTGTCTACAGGCTCAGCTCCACCCCTGTCTACAGAGCTCAGCTCTCTACAGCTTAA 1980
Dh 1468 CTGTCTACAGGCTCAGCTCCACCCCTGTCTACAGAGCTCAGCTCTCTACAGCTTAA 1527
Oy 1981 GTCCAGAGCCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2040
Dh 1528 GTCCAGAGCCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1587
Oy 2041 GGCAGTGAAGTACTTCTTGGGGGCAACAGACCCCTGATTTCTCTCACTCTACCCCTG 2100
Dh 1588 GGCAGTGAAGTACTTCTTGGGGGCAACAGACCCCTGATTTCTCTCACTCTACCCCTG 1647
Oy 2101 CACACTCAGCTGTGGCAGGGAATGAAGAGCTTCCGTCAGAAAAAGGCTCAGCCCTC 2160
Dh 1648 CACACTCAGCTGTGGCAGGGAATGAAGAGCTTCCGTCAGAAAAAGGCTCAGCCCTC 1707
Oy 2161 CCAGCCCGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2220
Dh 1708 CCAGCCCGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1767
Oy 2221 GGCAGGAGACAGTGGGCCCCCAAGAGGAGAGATCTCTGAGGCCGAGCCGAGGAGG 2280
Dh 1768 GGCAGGAGACAGTGGGCCCCCAAGAGGAGAGATCTCTGAGGCCGAGCCGAGGAGG 1827
Oy 2281 CTGGAAGGCTGGCAGATGCTTCCCTCATCTCCACCTCCAGCCGCTCAGGCTCTTCTCT 2340
Dh 1828 CTGGAAGGCTGGCAGATGCTTCCCTCATCTCCACCTCCAGCCGCTCAGGCTCTTCTCT 1887
Oy 2341 GTCCCAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2400
Dh 1888 GTCCCAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1947
Oy 2401 CTCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2460
Dh 1948 CTCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2007
Oy 2461 TGTACTGAGAATAAACAATTTTTCATATTTGTTTAAATAAAAAAAAAAAAAA 2517
Dh 2008 TGTACTGAGAATAAACAATTTTTCATATTTGTTTAAATAAAAAAAAAAAAAA 2064
```

RESULT 11
AX405828

```
LOCUS AX405828 2061 bp DNA linear PAT.14-JUN-2002
DEFINITION Sequence 243 from Patent WO222660.
ACCESSION AX405828
VERSION AX405828.1 GI:21439095
KEYWORDS
SOURCE
ORGANISM human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Tang,Y.T., Liu,C., Zhou,P., Asundi,V., Zhang,J., Zhao,Q.A., Ren,F.,
Xue,A.-J., Yang,Y., Wehrman,T. and Drmanac,R.T.
TITLE Novel nucleic acids and polypeptides
JOURNAL Patent: WO 022660-A 243 21-MAR-2002;
HYSEQ, INC. (US)
FEATURES
Source Location/Qualifiers
1..2061
/organism="Homo sapiens"
/db_xref="taxon:9606"
147..1331
CDS
/codon_start=1
/protein_id="CAD34935.1"
/db_xref="GI:21439096"
/translation="MALTVSRSPGSGASTPVGPMQAVORSHLORHROFAVLRLA
VLGLODGDNDDAEASSEPEKAPSEBELGDDTDFOGSOPKOEORHILMY
OLRPODDIRLAOLEAPRPRRLYLIVTRBEGLSODTLLVGFPPSSPCT
LGLVPLMSPTQVYLDGSGSVSGSRIFKPSISDTMMATLOLVQAEBALGSL
LPGSALVTMASSHIOERLNSFOSCLNENTAMADLESLEPSPAEKGSSEDDOMEALR
LELWKLVELESTSESDMPFVFSHESHEBLOFPLQALATKGGQVDRGPOPALKS
ROSYVTLTGSAVAVNRTQAFQEOEGOGGOGGECISSTPFRFRRVVRQASVHDSGEEG
EA"
BASE COUNT 415 a 672 c 605 g 369 t
ORIGIN
Query Match 47.9%; Score 1216.4; DB 6; Length 2061;
Best Local Similarity 74.6%; Pred. No. 2.3e-240;
Matches 1875; Conservative 0; Mismatches 126; Indels 513; Gaps 7;
Oy 1 CCTGTCTCTGGCGGTCGAGAGTGTCCCGGGGTTGAGAGAGGCGCTCCCGGTGC 60
Dh 61 CCTGTCTCTGGCGGTCGAGAGTGTCCCGGGGTTGAGAGAGGCGCTCCCGGTGC 119
Oy 61 CAGCCAGGTGTCTCGCGGCTGCTCATGAGCCCTGTGACAGTGAAGCCGTGCGCCCG 120
Dh 120 CAGCCAGGTGTCTCGCGGCTGCTCATGAGCCCTGTGACAGTGAAGCCGTGCGCCCG 179
Oy 121 GGCAGCGGCGCTTCCACGCGCTGGGGCCCTGGGACCAAGCGGTCCAGCAAGAGTCA 180
Dh 180 GGCAGCGGCGCTTCCACGCGCTGGGGCCCTGGGACCAAGCGGTCCAGCAAGAGTCA 239
Oy 181 CTCCAGCGAAGCAGAGCTTTTGGGTGCTTCCGTGGGGCTGTCTTGGAGTGAAGTGA 240
Dh 240 CTCCAGCGAAGCAGAGCTTTTGGGTGCTTCCGTGGGGCTGTCTTGGAGTGAAGTGA 299
Oy 241 GGGACAATGATGATGACAGCAGAGCCAGTTCTGAGCCAAAGAGAGGCCCGGAGTGA 300
Dh 300 GGGACAATGATGATGACAGCAGAGCCAGTTCTGAGCCAAAGAGAGGCCCGGAGTGA 359
Oy 301 GAGGAGCTCCACGGGGAGCCAGACAGACTTGGGCAAGAGATCCAGAGTCCCAAGAGAG 360
Dh 360 GAGGAGCTCCACGGGGAGCCAGACAGACTTGGGCAAGAGATCCAGAGTCCCAAGAGAG 419
Oy 361 GAGGAGCAGAGCAGCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
Dh 420 GAGGAGCAGAGCAGCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 479
Oy 421 GCGCTGGAGCCAGCTGGAGGACCCCGGCTCCCGGCTCGGTACTGTGGTGAAT 480
Dh 480 GCGCTGGAGCCAGCTGGAGGACCCCGGCTCCCGGCTCGGTACTGTGGTGAAT 539
Oy 481 TCTACACGAGAGAGAGTGTGAGCCAGAGTGAAGAGGCTCTCTGGGCGTGAATTC 540
```


REFERENCE 1
 AUTHORS Lucbe, R.M. and Wei, B.
 TITLE Dsp-15 dual-specificity phosphatase
 JOURNAL Patent: WO 02/4740-A 20 28-MAR-2002;
 Ceplyr, Inc. (US)
 FEATURES Location/Qualifiers
 source 1. 1416
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 BASE COUNT 288 a 450 c 449 g 229 t
 ORIGIN

Query Match 40.5%; Score 1028; DB 6; Length 1416;
 Best Local Similarity 86.7%; Pred. No. 1.6e-201;
 Matches 1227; Conservative 0; Mismatches 0; Indels 189; Gaps 1;

QY 88 ATGGCCCTGTCACAGTACAGCCGTTGCCCCGGGACAGGGCCCTCCACGCCCTGGGG 147
 DB 1 ATGGCCCTGTCACAGTACAGCCGTTGCCCCGGGACAGGGCCCTCCACGCCCTGGGG 60

QY 148 CCCGTGGACACAGGGGGTCCACAGAGAGTGCATCCAGCCGAAGGACAGCTTGGCGTG 207
 DB 61 CCCGTGGACACAGGGGGTCCACAGAGAGTGCATCCAGCCGAAGGACAGCTTGGCGTG 120

QY 208 CTCCTGGGGGCTGCTCTGGAGCTCAGAGTGGAGGGACATGATGATGACAGAGGCC 267
 DB 121 CTCCTGGGGGCTGCTCTGGAGCTCAGAGTGGAGGGACATGATGATGACAGAGGCC 180

QY 268 AGTTCTGAGCCACAGAGAGAGGCCCGCAGTGGAGAGAGCTCCAGGGGACAGACAGAC 327
 DB 181 AGTTCTGAGCCACAGAGAGAGGCCCGCAGTGGAGAGAGCTCCAGGGGACAGACAGAC 240

QY 328 TTGCGGACAGAGATCCCGAGAGTCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 387
 DB 241 TTGCGGACAGAGATCCCGAGAGTCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300

QY 388 ATGATACAGCTGTCTGAGGCGCAGAGATGACATCCGCTGGAGGCCAGCTGGAGCACCC 447
 DB 301 ATGATACAGCTGTCTGAGGCGCAGAGATGACATCCGCTGGAGGCCAGCTGGAGCACCC 360

QY 448 CGGCTCCCGCGGCTCCGCTACCTGCTGTTCTCTACACAGAGAGAGAGAGTCTGAGC 507
 DB 361 CGGCTCCCGCGGCTCCGCTACCTGCTGTTCTCTACACAGAGAGAGAGAGTCTGAGC 420

QY 508 CAGATGAGAGAGGTCCTCTGGGGGCTGGATTTCCCTGACACAGCTCCCGCAGCTGCACC 567
 DB 421 CAGATGAGAGAGGTCCTCTGGGGGCTGGATTTCCCTGACACAGCTCCCGCAGCTGCACC 480

QY 568 CTGGGCGCTGCTTGGCCCTCTGGAGTACACAGAGTGTACTTATGATGAGAGGGGGC 627
 DB 481 CTGGGCGCTGCTTGGCCCTCTGGAGTACACAGAGTGTACTTATGATGAGAGGGGGC 540

QY 628 TTTCAGCGTACGTCGTGTGGGCAAGCCGGATCTTCAAGCCCATCTCCATCCAGACATG 687
 DB 541 TTTCAGCGTACGTCGTGTGGGCAAGCCGGATCTTCAAGCCCATCTCCATCCAGACATG 600

QY 688 T----- 688
 DB 601 TGGCCACACTCCAGGATTTGACCAAGCATGTAGAGCAGCTTAAGGACGCGCTTGTGA 660

QY 689 ----- 688
 DB 661 CCGGGTGGCAATGGCCCTCAGCTGGGCCAGCCACTACCAAGAGAGACTGAATCCGAACAG 720

QY 689 ----- 688
 DB 721 AGTCCCTCAATGATGAGAGCGCTATGGCCGACCTGAGTCTCTGGCGCTCCAGCGCC 780

QY 689 ----- 738
 DB 781 GAGCCTGGCGGGTCTCTGAAACAGAGAGAGATGAGAGCGGCGATCCGCTGAGCTGTGG 840

QY 739 AAAGTGTGATGTACAGTACCTGAGAGTGTCACTTCCAAAGAGATCCGACAGGCTCTG 798

DB 841 AAAGTGTGATGTACAGTACCTGAGAGTGTCACTTCCAAAGAGATCCGACAGGCTCTG 900
 QY 799 GAGCTGGCCCTGGGGCTCTCCCTCCAGAGTACCGTACTTATCATGACACCAATGCTG 858
 DB 901 GAGCTGGCCCTGGGGCTCTCCCTCCAGAGTACCGTACTTATCATGACACCAATGCTG 960
 QY 859 CTGCTGGTGGACAGCGGAGCGGAGCCCTCCGCAATCTTCCCGACCTTACTGGGCTCA 918
 DB 961 CTGCTGGTGGACAGCGGAGCGGAGCCCTCCGCAATCTTCCCGACCTTACTGGGCTCA 1020

QY 919 GAGTGAACGACGAAACCTTGAGAGGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAG 978
 DB 1021 GAGTGAACGACGAAACCTTGAGAGGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080

QY 979 ATGGCCCGGGAGATTGACACTTTCACCTGAGCGCTTACCTACCAATGTCGCTC 1038
 DB 1081 ATGGCCCGGGAGATTGACACTTTCACCTGAGCGCTTACCTACCAATGTCGCTC 1140

QY 1039 TGGATGAGAGTGGGCCACCTGCTGCGCAGTGGAGAGAGAGAGAGAGAGAGAGAGAG 1098
 DB 1141 TGGATGAGAGTGGGCCACCTGCTGCGCAGTGGAGAGAGAGAGAGAGAGAGAGAGAG 1200

QY 1099 GCTGCAAGACGACAGGGGACCCAGCTGCTGCTCAGTCAAGATGAGCGCTCAGCGCTCA 1158
 DB 1201 GCTGCAAGACGACAGGGGACCCAGCTGCTGCTCAGTCAAGATGAGCGCTCAGCGCTCA 1260

QY 1159 GCGGCGACAGTGTGCGCTATGCAATGAGAGAGTACGATGAGAGAGAGAGAGAGAG 1218
 DB 1261 GCGGCGACAGTGTGCGCTATGCAATGAGAGAGTACGATGAGAGAGAGAGAGAGAG 1320

QY 1219 CGCCACGTGACAGAGCTCCGGCCCATCGCCGCCCAACCTGCTTCCTGCGCAGCTG 1278
 DB 1321 CGCCACGTGACAGAGCTCCGGCCCATCGCCGCCCAACCTGCTTCCTGCGCAGCTG 1380

QY 1279 CAGATCTACAGGCGCATCTGACGCGCCAGAACCTGA 1314
 DB 1381 CAGATCTACAGGCGCATCTGACGCGCCAGAACCTGA 1416

RESULT 13
 AP002776 160903 bp DNA linear HTG 17-JAN-2001
 LOCUS Homo sapiens chromosome 11 clone RP11-126P21 map 11q, WORKING DRAFT
 DEFINITION
 SEQUENCE, 36 unordered pieces.
 AP002776 GI:12246853
 VERSION
 HTG: HTGS-PHASE1; HTGS-DRAFT.
 KEYWORDS
 Homo sapiens DNA, clone:RP11-126P21.
 SOURCE
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 (bases 1 to 160903)
 Hattori, M., Ishii, K., Toyoda, A., Taylor, T. D., Hong-Seong, P.,
 Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
 Homo sapiens 160,903 genomic DNA of 11q
 Published only in Database (2000)
 2 (bases 1 to 160903)
 Hattori, M., Ishii, K., Toyoda, A., Taylor, T. D., Hong-Seong, P.,
 Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
 Direct Submission
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 Submitted (10-JUL-2000) Masahira Hattori, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 (E-mail: hattori@egsc.riken.go.jp, URL: http://hnp.gsc.riken.go.jp/,
 Tel: 81-45-503-9111, Fax: 81-45-503-9170)
 On Jan 16, 2001 this sequence version replaced gi:1188614.
 ----- Genome Center
 Center: RIKEN Genomic Sciences Center(GSC)
 Center code: RIKEN
 Web site: http://hnp.gsc.riken.go.jp/
 Contact: hattori@egsc.riken.go.jp
 ----- Project Information

are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

```
1 15114 contig of 15114 bp in length
15215 27885 contig of 12671 bp in length
27986 42419 contig of 14434 bp in length
42520 55000 contig of 12481 bp in length
55101 68523 contig of 13423 bp in length
68624 79107 contig of 10484 bp in length
79208 89207 contig of 10000 bp in length
89308 95794 contig of 6487 bp in length
95895 103725 contig of 7831 bp in length
103826 112105 contig of 8280 bp in length
112206 119213 contig of 7008 bp in length
119314 124281 contig of 4968 bp in length
124382 130027 contig of 5646 bp in length
130128 135012 contig of 4885 bp in length
135113 138400 contig of 3888 bp in length
138501 144921 contig of 6421 bp in length
144922 150480 contig of 5459 bp in length
150481 150580 contig of 5459 bp in length
150581 153837 contig of 3011 bp in length
153838 153937 contig of 3257 bp in length
153938 156948 contig of 3011 bp in length
156949 157048 gap of 100 bp
```

* NOTE: This is a 'working draft' sequence. It currently consists of 24 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```
1 15114 contig of 15114 bp in length
15115 15214 gap of 100 bp
15215 27885 contig of 12671 bp in length
27886 27985 gap of 100 bp
27986 42419 contig of 14434 bp in length
42420 42519 gap of 100 bp
42520 55000 contig of 12481 bp in length
55001 55100 gap of 100 bp
55101 68523 contig of 13423 bp in length
68524 68623 gap of 100 bp
68624 79107 contig of 10484 bp in length
79108 79207 gap of 100 bp
79208 89207 contig of 10000 bp in length
89208 89307 gap of 100 bp
89308 95794 contig of 6487 bp in length
95795 95894 gap of 100 bp
95895 103725 contig of 7831 bp in length
103726 103825 gap of 100 bp
103826 112105 contig of 8280 bp in length
112106 112205 gap of 100 bp
112206 119213 contig of 7008 bp in length
119214 119313 gap of 100 bp
119314 124281 contig of 4968 bp in length
124282 124381 gap of 100 bp
124382 130027 contig of 5646 bp in length
130028 130127 gap of 100 bp
130128 135012 contig of 4885 bp in length
135013 135112 gap of 100 bp
135113 138400 contig of 3888 bp in length
138401 138500 gap of 100 bp
138501 144921 contig of 6421 bp in length
144922 145021 gap of 100 bp
145022 150480 contig of 5459 bp in length
150481 150580 gap of 100 bp
150581 153837 contig of 3011 bp in length
153838 153937 gap of 100 bp
153938 156948 contig of 3011 bp in length
156949 157048 gap of 100 bp
```

FEATURES

Source

1. 15114

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="11"

/map="11q"

/clone="RP11-157K17"

1. 15114

/note="assembly-fragment clone_end:T7 vector_side:right"

15215. 27885

/note="assembly-fragment"

27986. 42419

/note="assembly-fragment"

42520. 55000

/note="assembly-fragment"

55101. 68523

/note="assembly-fragment"

68624. 79107

/note="assembly-fragment"

79208. 89207

/note="assembly-fragment"

89308. 95794

/note="assembly-fragment"

95895. 103725

/note="assembly-fragment"

103826. 112105

/note="assembly-fragment"

112206. 119213

/note="assembly-fragment"

119314. 124281

/note="assembly-fragment"

124382. 130027

/note="assembly-fragment"

130128. 135012

/note="assembly-fragment"

135113. 138400

/note="assembly-fragment"

138501. 144921

/note="assembly-fragment"

145022. 150480

/note="assembly-fragment"

150581. 153837

/note="assembly-fragment"

153938. 156948

/note="assembly-fragment"

157049. 161384

/note="assembly-fragment"

161485. 164413

/note="assembly-fragment"

164514. 167396

/note="assembly-fragment"

167497. 169991

/note="assembly-fragment"

170092. 171566

/note="assembly-fragment"

170092. 171566

/note="assembly-fragment"

Location/Qualifiers

1. 171566

157049 161384: contig of 4336 bp in length

161385 161484: gap of 100 bp

161485 164413: contig of 2929 bp in length

164414 164513: gap of 100 bp

164514 167396: contig of 2883 bp in length

167397 167496: gap of 100 bp

167497 169991: contig of 2495 bp in length

169992 170091: gap of 100 bp

170092 171566: contig of 1475 bp in length.

BASE COUNT 45656 a 39885 c 40004 g 43721 t 2300 others

ORIGIN

Query Match 39.88; Score 1012; DB 2; Length 171566;
Best Local Similarity 100.0%; Pred. No. 5.4e-198;
Matches 1012; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1486 AGGCTCTCTCCAGAGCTTCACATGAGAGGCTTCGACAGCTTCACAGCTTG 1545
|||||

[illegible]

RESULT 15	AB072360	1416 bp	linear	PRI 29-JAN-2002
LOCUS	AB072360			
DEFINITION	Homo sapiens mRNA for hssh-3, complete cds.			
ACCESSION	AB072360			
VERSION	AB072360.1	GI:18376668		

KEYWORDS	source	score	length	gaps
ORGANISM	Homo sapiens cdna to mRNA.			
REFERENCE	Homo sapiens			
TITLE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
JOURNAL	Niwa, R., Nagata-Ohashi, K., Takeichi, M., Mizuno, K. and Uemura, T.			
MEDLINE	Control of actin reorganization by Slingshot, a family of phosphatases that dephosphorylate ADP/cofilin			
AUTHORS	Cell 108 (2), 233-246 (2002)			
JOURNAL	2 (bases 1 to 1416)			
MEDLINE	Niwa, R., Nagata-Ohashi, K., Hay, B.A., Takeichi, M., Mizuno, K. and Uemura, T.			
AUTHORS	Direct Submision			
TITLE	Submitted (29-SEP-2001) Tadashi Uemura, Laboratory of Molecular Genetics, The Institute for Virus Research, Kyoto University;			
JOURNAL	Shogoin-kawata-cho 53, Sakyo-ku, Kyoto 606-8507, Japan			
MEDLINE	(E-mail: tuemura@virus.kyoto-u.ac.jp, Tel:81-751-4031, Fax:81-751-751-3989)			
AUTHORS	Location/Qualifiers			
FEATURES	1..1416			
source	/organism="Homo sapiens"			
gene	/db_xref="taxon:9606"			
CDS	1..1416			
	/gene="hSSH-3"			
	1..1416			
	/gene="hSSH-3"			
	/standard_name="slingshot-3"			
	/note="a putative dual specific phosphatase"			
	/codon_start=1			
	/product="hSSH-3"			
	/protein_id="BAB84119.1"			
	/db_xref="GI:18376669"			
	/translation="MALVTVSRSPSGASTPYGPMQVQVQSRRLQROSFVLRGA			
	VLGIQSGDNDDAVEASSEPTERKPSSEELHGDQTDGQSSQSPQKEEDRHLHMY			
	QLEPRQDRIIAQLAQLERPRPRRLYLVLVSTREEGSLQEVVLGVDFPDSSPSCY			
	LGVLPLMSDTQVYLDIDGGSVYSGQSSIFRPISSIQTMATLVYLAQCEALGSG			
	LVPGSALITWASHYQERLNSEQSLINEMTAMADLESIRPSAEPGSSDEQEAOLR			
	ALMKVLDVDSLESVTSKEIRQLMSKAGMSPSSVTVSSPTTBCQWMSGDRASRIEP			
	HLTIGSEWMNANLEIQLORNYTHILNMARLIDNFPRTFYHNVRIMLDESSAOLPRW			
	KETHRTTEAARQGTHTVLCRKMVSRSAATVLAIVAMKQYESLQALRHVQELRPIA			
	RPNPGEFRLQIYQGLIARPT"			
BASE COUNT	288 a 449 c 449 g 230 t			
ORIGIN				
Query Match	39.5%; Score 1003.4; DB 9; Length 1416;			
Best Local Similarity	86.5%; Pred. No. 1.9e-196;			
Matches 1225; Conservative	0; Mismatches 1; Indels 191; Gaps 3;			
QY	88 ATGGCCCTGTGTACAGTGTAGCGCTTGCCGCCCGGACAGCGGCGCTCCAGCCCGGTGGG 147			
DB	1 ATGGCCCTGTGTACAGTGTAGCGCTTGCCGCCCGGACAGCGGCGCTCCAGCCCGGTGGG 60			
QY	148 CCCTGGGACCAAGCGCGGTCCAGCGAAGAGTGTACTCCAGCAAGGCAAGAGCTTTGGCGTG 207			
DB	61 CCCTGGGACCAAGCGCGGTCCAGCGAAGAGTGTACTCCAGCAAGGCAAGAGCTTTGGCGTG 120			
QY	208 CTCCTGGGCGCTGTCTGTGGACTGTGAGATGAGGGGCAATGATGATGACGACAGAGCC 267			
DB	121 CTCCTGTGGGCGCTGTCTGTGGACTGTGAGATGAGGGGCAATGATGATGATGAGAGGCC 180			
QY	268 AGTTGTGAGCCCAAGCAAGAGGCCCGCGAGTGAAGAGAGGTCCACGGGGGACACACAGAC 327			
DB	161 AGTTGTGAGCCCAAGCAAGAGGCCCGCGAGTGAAGAGAGGTCTCACGGGGGACACAGACAC 240			
QY	328 TTTCGGCGAAGATCCACAGTCCCGCAGAGAGAGAGAGAGAGAGGACACACTCCACTC 387			
DB	241 TTTCGGCGAAGATCCCGAAGTCCCGCAGAGAGAGAGAGAGAGAGGACACACTCCACTC 300			
QY	388 ATGTACAGACTGTCTGAGCGCGAGAGTATCCGCTTGCGAGCCCACTGGAGGACACC 447			
DB	301 ATGTACAGACTGTCTGAGCGCGAGAGTATCCGCTTGCGAGCCCACTGGAGGACACC 360			

XX (PEKE) PE CORP NY.
 XX
 XX
 PI Wei M, Ketchum KA, Di Francesco V, Beasley EM;
 XX
 DR WPI: 2002-575237/61.
 DR P-PSDB; ABP51655.
 DR
 XX
 XX Novel isolated human phosphatase peptide useful for treating disorder
 PT characterized by absence of, inappropriate or unwanted expression of
 PT the phosphatase protein, and as immunogens to raise antibodies -
 XX
 PS Claim 1; Fig 1C; 85pp; English.

CC The present invention describes an isolated human phosphatase peptide
CC (1). (1) can be used for identifying a modulator of (1) by contacting
CC (1) with an agent and determining if the agent has modulated the
CC function or activity of (1). (1) is useful for identifying an agent that
CC binds to (1), by contacting (1) with an agent and assaying the contacted
CC mixture to determine whether a complex is formed with the agent bound
CC (1). The human phosphatases from the present invention are mitogen
CC activated protein (MAP) kinase phosphatases. These human MAP kinase
CC phosphatases are located on chromosome 11. (1) and the polynucleotide
CC sequences encoding (1) can be used in gene therapy. The present sequence
CC encodes human MAP kinase phosphatase splice form 3 from the present
CC invention.
XX
XX Sequence 2540 BP; 552 A; 809 C; 739 G; 440 T; 0 other;
XX

Query Match	100.0%	Score 2540;	DB 24;	Length 2540;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2540; Conservative	0;	Mismatches	0;	Gaps 0

[illegible]

QY	601	CAGGTGACTTAGATGATGGAGACGGGGGCTTCAGCTGACGTGAGTGGGGCAAAAGCCGATC	660
Db	601	CAGGTGACTTAGATGATGGAGACGGGGGCTTCAGCTGACGTGAGTGGGGCAAAAGCCGATC	660
QY	661	TTCAAGCCATCTCCATCCAGACCATGTGTCTCTCGAAGACAGAGACAGATGGAGCAGCG	720
Db	661	TTCAAGCCATCTCCATCCAGACCATGTGTCTCTCGAAGACAGAGACAGATGGAGCAGCG	720
QY	721	ATCCGTGCTGAGCTGTGTGGAAAGTTGTGTGATGTAGTGACCTTGACAGTGTCACTTCCAA	780
Db	721	ATCCGTGCTGAGCTGTGTGGAAAGTTGTGTGATGTAGTGACCTTGACAGTGTCACTTCCAA	780
QY	781	GAGATCCGCCAGGGCTGTGGAGCTGCGGCTTGGGGCTCCCTCCACAGACAGCTGACTTC	840
Db	781	GAGATCCGCCAGGGCTGTGGAGCTGCGGCTTGGGGCTCCCTCCACAGACAGCTGACTTC	840
QY	841	ATCGACAACAGATGCTGTCTGTGTGTGGACAGCGGGACCGAGCTCCCGCATCTTCCCC	900
Db	841	ATCGACAACAGATGCTGTCTGTGTGTGGACAGCGGGACCGAGCTCCCGCATCTTCCCC	900
QY	901	CACCTTACCTGGGGCTCGAGATGGGAACCGCAACCTGGAGAGACTGAGAGGAACAGG	960
Db	901	CACCTTACCTGGGGCTCGAGATGGGAACCGCAACCTGGAGAGACTGAGAGGAACAGG	960
QY	961	GTCACCCACATCTTGAACATGTGGCCCGGGAGATTGACAACTTGTACCTGTAGGCGCTTACC	1020
Db	961	GTCACCCACATCTTGAACATGTGGCCCGGGAGATTGACAACTTGTACCTGTAGGCGCTTACC	1020
QY	1021	TACCAACATGTGGCCCTGTGGATGAGGAGTGGGGCTGGACCTGTGCGGACTGTGGAGGAG	1080
Db	1021	TACCAACATGTGGCCCTGTGGATGAGGAGTGGGGCTGGACCTGTGCGGACTGTGGAGGAG	1080
QY	1081	ACGCACCGCTTCATTGAGGCTGTGAAGAAGCAAGGAGACCCACGTCGTGTCTCACTGCAAG	1140
Db	1081	ACGCACCGCTTCATTGAGGCTGTGAAGAAGCAAGGAGACCCACGTCGTGTCTCACTGCAAG	1140
QY	1141	ATGGGCGTCACCGGCTGTAGGGGGCACAGTGTGGCTTAGTCATGAAGAGTACGAATGC	1200
Db	1141	ATGGGCGTCACCGGCTGTAGGGGGCACAGTGTGGCTTAGTCATGAAGAGTACGAATGC	1200
QY	1201	AGCCTGGAAGGCGCTGTGCGCCACAGTGTCAAGAGACCTCGGCGCCATGCGCCGCCAACCT	1260
Db	1201	AGCCTGGAAGGCGCTGTGCGCCACAGTGTCAAGAGACCTCGGCGCCATGCGCCGCCAACCT	1260
QY	1261	GCGTTCCTGCGCCAGCTGTGACATCTACCAAGGACTCTCTACAGGCCCAAGACCTGAGGGTGG	1320
Db	1261	GCGTTCCTGCGCCAGCTGTGACATCTACCAAGGACTCTCTACAGGCCCAAGACCTGAGGGTGG	1320
QY	1321	TGGGAGGAGAGAAGTTGTAGGCATGTGAAGAGACCAGCGAGCCCGAAAGAGAACCTGG	1380
Db	1321	TGGGAGGAGAGAAGTTGTAGGCATGTGAAGAGACCAGCGAGCCCGAAAGAGAACCTGG	1380
QY	1381	GGCCACGGGSCAGCTATPAACCTCCGAGGGGTCATGAGTCCATCAGTCTTCTGTGAGCC	1440
Db	1381	GGCCACGGGSCAGCTATPAACCTCCGAGGGGTCATGAGTCCATCAGTCTTCTGTGAGCC	1440
QY	1441	TGCTTGGAGCTGTGAGAGCACTCAGTAGAGACAGTGAACATGCGAAGAGGTCCTTCTCC	1500
Db	1441	TGCTTGGAGCTGTGAGAGCACTCAGTAGAGACAGTGAACATGCGAAGAGGTCCTTCTCC	1500
QY	1501	ACGAGTCTTCACATGAGAGAGCTTGTGAGCGCTTCCACAGGTTGCAAGGACCAAGGAG	1560
Db	1501	ACGAGTCTTCACATGAGAGAGCTTGTGAGCGCTTCCACAGGTTGCAAGGACCAAGGAG	1560
QY	1561	GCCACAGGTGGACAGGGGGCTTGAGCTTGCCTGAATGCCGCAAGTCAGTAGTGTACCC	1620
Db	1561	GCCACAGGTGGACAGGGGGCTTGAGCTTGCCTGAATGCCGCAAGTCAGTAGTGTACCC	1620
QY	1621	TTCACAGGAGAGGCGTGTGTGGCCAACCGGACCCAGGCGCTTCAGAGGAGCAGGAGCGGC	1680
Db	1621	TTCACAGGAGAGGCGTGTGTGGCCAACCGGACCCAGGCGCTTCAGAGGAGCAGGAGCGGC	1680
QY	1681	AGGGGACAGGGGACAGGAGAGCCCTGTCTCTACGCGCCAGAGTCTCCGAGAGTGTGTGA	1740

```

Db 1661 AGGGGAGGGGAGGAGAGGCGCTGCACTTTCTCTACGCCAGGTTCCGGAGGCTGTGA 1740
Qy 1741 GACAGAGCCAGCGTGCATGACAGTGGAGAGGAGGCGGAGCCCGAGCCGTCACATGCCC 1800
Db 1741 GACAGAGCCAGCGTGCATGACAGTGGAGAGGAGGCGGAGCCCGAGCCGTCACATGCCC 1800
Qy 1801 ACAGTCCCGCTGACACTGAAAGAGATCAACAACCTCTTGGAGAAACCCCTCAGCTGTG 1860
Db 1801 ACAGTCCCGCTGACACTGAAAGAGATCAACAACCTCTTGGAGAAACCCCTCAGCTGTG 1860
Qy 1861 GCCGACACATTTCTCTGAGTCCGCCCCATACCCGTCACAGCCCTACCTCCACC 1920
Db 1861 GCCGACACATTTCTCTGAGTCCGCCCCATACCCGTCACAGCCCTACCTCCACC 1920
Qy 1921 CTGTACAGAGGCGCTCAGCTCCGACCCGCTGACAGCCCTACCTCCCTACAGCTTAA 1980
Db 1921 CTGTACAGAGGCGCTCAGCTCCGACCCGCTGACAGCCCTACCTCCCTACAGCTTAA 1980
Qy 1981 GTCCAGAGCCATGTCTGCTGTCCAGAGGCTCAAGACTTTCTAAGTGGATGTGAGA 2040
Db 1981 GTCCAGAGCCATGTCTGCTGTCCAGAGGCTCAAGACTTTCTAAGTGGATGTGAGA 2040
Qy 2041 GGGAGTGAAGGTACCTTTGGGGGCAACAGACACCTAGTTTCAATCTCAGCTAGCCCTG 2100
Db 2041 GGGAGTGAAGGTACCTTTGGGGGCAACAGACACCTAGTTTCAATCTCAGCTAGCCCTG 2100
Qy 2101 CACACTCAGCTGTGGCAGAGAAATGAAAACAGAGCTCCCGTCAAAAAAGGTCACGCCCTC 2160
Db 2101 CACACTCAGCTGTGGCAGAGAAATGAAAACAGAGCTCCCGTCAAAAAAGGTCACGCCCTC 2160
Qy 2161 CCACCCCGCCGCTCCTCTGACACCTCTCTCTCCAGTTTCAATCTCGAAGCAGCA 2220
Db 2161 CCACCCCGCCGCTCCTCTGACACCTCTCTCTCCAGTTTCAATCTCGAAGCAGCA 2220
Qy 2221 GGGCAGGACCAAGTGGCCCCCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2280
Db 2221 GGGCAGGACCAAGTGGCCCCCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2280
Qy 2281 CTGGAAGGGGCGGAGATGCTTCCCTCATGCACTCCACCGGTCAGAGTCTTCTGCTCT 2340
Db 2281 CTGGAAGGGGCGGAGATGCTTCCCTCATGCACTCCACCGGTCAGAGTCTTCTGCTCT 2340
Qy 2341 GTCCCAAGCTCTCTGTGACACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2400
Db 2341 GTCCCAAGCTCTCTGTGACACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2400
Qy 2401 CTTTGTGCTCTTCTGCGCTGCTGCTAGTCAAGTTTTCATAGCCTTACAGTATCTGCTT 2460
Db 2401 CTTTGTGCTCTTCTGCGCTGCTGCTAGTCAAGTTTTCATAGCCTTACAGTATCTGCTT 2460
Qy 2461 TGTACTGGAATTAACACATTTTTCATATTTGTTAAAAAAGGAGGAGGAGGAGGAGGAGG 2520
Db 2461 TGTACTGGAATTAACACATTTTTCATATTTGTTAAAAAAGGAGGAGGAGGAGGAGGAGG 2520
Qy 2521 AAAAAAAAAAAAAAAAAA 2540
Db 2521 AAAAAAAAAAAAAAAAAA 2540

```

RESULT 2
AB073249
ID AB073249 standard; cDNA; 2704 BP.

AC AB073249;
XX 30-SEP-2002 (first entry)
DE Human MAP kinase phosphatase splice form 1 cDNA sequence SEQ ID NO:1.
XX Human: phosphatase; mitogen activated protein kinase phosphatase;
KW MAP kinase; enzyme; chromosome 11; gene; SS.
XX

```

OS Homo sapiens.
XX Key Location/Qualifiers
FX 5'UTR 1..93
FT /tag= a
FT CDS 94..1509
FT /tag= b
FT /product= "MAP kinase phosphatase splice form 1"
FT 3'UTR 1510..2704
FT /tag= c
XX
XX WO200242436-A2.
XX 30-MAY-2002.
XX 07-NOV-2001; 2001WO-US42995.
XX 20-NOV-2000; 2000US-0715177.
XX 18-JAN-2001; 2001US-0761640.
XX (PEKE ) PE CORP NY.
XX
XX Wei M, Ketchum KA, Di Francesco V, Beasley EM;
XX WPI; 2002-575237/61.
XX P-PSDB; ABP51653.
XX
XX Novel isolated human phosphatase peptide useful for treating disorder
XX characterized by absence of, inappropriate or unwanted expression of
XX the phosphatase protein, and as immunogens to raise antibodies
XX
XX Claim 1; Fig 1A; 85pp; English.
XX
XX The present invention describes an isolated human phosphatase peptide
XX (I). (I) can be used for identifying a modulator of (I) by contacting
XX (I) with an agent and determining if the agent has modulated the
XX function or activity of (I). (I) is useful for identifying an agent that
XX binds to (I), by contacting (I) with an agent and assaying the contacted
XX mixture to determine whether a complex is formed with the agent bound
XX (I). The human phosphatases from the present invention are mitogen
XX activated protein (MAP) kinase phosphatases. These human MAP kinase
XX phosphatases are located on chromosome 11. (I) and the polynucleotide
XX sequences encoding (I) can be used in gene therapy. The present sequence
XX encodes human MAP kinase phosphatase splice form 1 from the present
XX invention.
XX
XX Sequence 2704 BP; 569 A; 874 C; 794 G; 467 T; 0 other.
XX
XX Query Match 88.4%; Score 2246.4; DB 24; Length 2704;
XX Best Local Similarity 92.4%; Pred. No. 0;
XX Matches 2497; Conservative 0; Mismatches 11; Indels 194; Gaps 5;
Qy 1 CTTGTCTCTGCGGTCAGAGACTGTCGCGGGGTTGAAGGAGGAGGAGGAGGAGGAGGAGG 60
Db 8 CTTGTCTCTGCGGTCAGAGACTGTCGCGGGGTTGAAGGAGGAGGAGGAGGAGGAGGAGG 66
Qy 61 CAGCCAGAGTGTCTGCGGCGCTCCATGAGCCCTGTGACAGTATAGCCGTTCCGCCCGC 120
Db 67 CAGCCAGAGTGTCTGCGGCGCTCCATGAGCCCTGTGACAGTATAGCCGTTCCGCCCGC 126
Qy 121 GGCAGCGGCGCTCCAGCCCGCTGGGGCCCTGGGACACAGGCGGTCCAGCAAGAGTGA 180
Db 127 GGCAGCGGCGCTCCAGCCCGCTGGGGCCCTGGGACACAGGCGGTCCAGCAAGAGTGA 186
Qy 181 CTCAGCGCAAGGACAGCTTTGCGGTCTCGTGGGCTGTCTGAGACTGCAGAGTGA 240
Db 187 CTCAGCGCAAGGACAGCTTTGCGGTCTCGTGGGCTGTCTGAGACTGCAGAGTGA 246
Qy 241 GGGGACATGATGATGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
Db 247 GGGGACATGATGATGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 306
Qy 301 GAGAGCTCCAGCGGAGCAGACAGACTTTCGGGCAAGAGATCCAGAGTCCCGAGAGCAG 360

```

Db	1387	TACAAATGACAGCCGTGAGCAGAGCCCTGGCCGCAACGTGTGACAGAACTCCGGCCCATGCCCCG	1446
QY	1252	CCCAACCCCTGGCTTCCTGGCCGCAAGCTGCGAGATCTACAGAGGCATCTAGAGCCGCAAAAC	1311
Db	1447	CCCAACCCCTGGCTTCCTGGCCGCAAGCTGCGAGATCTACAGAGGCATCTAGAGCCGCAAAAC	1506
QY	1312	TGAGGCTGTGGGGAGAGAAAGTTGTATGATATGGAAGAAGACCAGAGCCCGGAAGA	1371
Db	1507	TGAGGGGTGGGGAGAGAAAGTTGTATGATATGGAAGAAGACCAGAGCCCGGAAGA	1566
QY	1372	AGAGCCCTGGGGGCAAGGAGCAGGTATTAACCTCCAGAGGGTCAATAGAGTCAATAGCTT	1431
Db	1567	AGAGGCTGGG - CCAGGCGCCACGATTAATAACCTCCAGAGGGTCAATAGAGTCAATAGCTT	1624
QY	1432	CTGAGCCCTCCTTTGGGAGCTTGAAGAGCACTAGTATAGACAGAGATACGACGAGAGTCT	1491
Db	1625	CTGAGCCCTCCTTT - GGAAGCTTGAAGAGCACTAGTATAGACAGAGATACGAGAGTCT	1682
QY	1492	TCCTCTCCACAGAGTCTTCAATGAAGAGCTCTGCAAGCCCTTCCCAACAGCTTGCAGAAGA	1551
Db	1683	TCCTCTCCACAGAGTCTTCAATGAAGAGCTCTGCAAGCCCTTCCCAACAGCTTGCAGAAGA	1742
QY	1552	CCAGAGGAGCCACGACGATGAGACAGAGGGGCTCAGCTGCCTCGGAAGTCCCGCAATCAG	1611
Db	1743	CCAGAGGAGCCACGACGATGAGACAGAGGGGCTCAGCTGCCTCGGAAGTCCCGCAATCAG	1802
QY	1612	TGGTTACCTCCAGGCGCGAGTCCGCTGGTGGCCAAACCGGACCCAGGCGCTTCCAGAGACAG	1671
Db	1803	TGGTTACCTCCAGGCGCGAGTCCGCTGGTGGCCAAACCGGACCCAGGCGCTTCCAGAGACAG	1862
QY	1672	AGCAGGGGACAGGGGACAGGGGACAGGAGAGGCCGCAATTTCTCTACAGCCAGGTTCCGGA	1731
Db	1863	AGCAGGGGACAGGGGACAGGGGACAGGAGAGGCCGCAATTTCTCTACAGCCAGGTTCCGGA	1922
QY	1732	AGGAGGTAGACAGGCGCAGGCTGCATGACAGTGTGAGAGGAGGCGCAGGCTGAGCCCTCA	1791
Db	1923	AGGAGGTAGACAGGCGCAGGCTGCATGACAGTGTGAGAGGAGGCGCAGGCTGAGCCCTCA	1982
QY	1792	CACATGCCACAGCTCCCTGAGACACTGAAGAGATCCCAACTCTCTTGAAGAAACACCTC	1851
Db	1983	CACATGCCACAGCTCCCTGAGACACTGAAGAGATCCCAACTCTCTTGAAGAAACACCTC	2042
QY	1852	ACGCTCTTCCCGCACACAATTCCTCTCAGCTCCGCCCATACCCGTCACACTACAGCCTCAC	1911
Db	2043	ACGCTCTTCCCGCACACAATTCCTCTCAGCTCCGCCCATACCCGTCACACTACAGCCTCAC	2102
QY	1912	CTCCACCCCTGTCACTACAGGCGCTACCTCCACCCCTGTACATACAGCCGTCACACTCTCA	1971
Db	2103	CTCCACCCCTGTCACTACAGGCGCTACCTCCACCCCTGTACATACAGCCGTCACACTCTCA	2162
QY	1972	CACCTTAAAGTCCCAAGGCGCAGTGTGCTGTCCAAAGGGCTCAAGACTTCTTAACTGGGA	2031
Db	2163	CACCTTAAAGTCCCAAGGCGCAGTGTGCTGTCCAAAGGGCTCAAGACTTCTTAACTGGGA	2222
QY	2032	TGTGGTGAAGGAGTGAAGATACCTTTGGGGCAACAGCAACCCAGTTTCATTTCTCAACT	2091
Db	2223	TGTGGTGAAGGAGTGAAGATACCTTTGGGGCAACAGCAACCCAGTTTCATTTCTCAACT	2282
QY	2092	CTAGCCCTGCAACACTCAGCTGTGGCAAGGATGAAGACAGAGCTTCCCGTCAAAAAAGGG	2151
Db	2283	CTAGCCCTGCAACACTCAGCTGTGGCAAGGATGAAGACAGAGCTTCCCGTCAAAAAAGGG	2342
QY	2152	TCAGCCTTCCACCCCGCCCTCCCTGCAACCTCCGCTCTCTCCAGTTTATTTCTGG	2211
Db	2343	TCAGCCTTCCACCCCGCCCTCCCTGCAACCTCCGCTCTCTCCAGTTTATTTCTGG	2402
QY	2212	AACCAAGCCAGGCGCAGGCAACCAAGTGGCCCCCAAGGCAAGGAGATCTCCAGGCCCAAGC	2271
Db	2403	AACCAAGCCAGGCGCAGGCAACCAAGTGGCCCCCAAGGCAAGGAGATCTCCAGGCCCAAGC	2462
QY	2272	CGCGGAGGCTGGAAGAGGCTGTGAGATATGCTTCCCTCATACCACTCCACCGGTCCAGGTC	2331
Db	2463	CGCGGAGGCTGGAAGAGGCTGTGAGATATGCTTCCCTCATACCACTCCACCGGTCCAGGTC	2522

Oy	2332	TTTCTGCTGTGCCCGACAGACTCCTGTGACACGACGAGATCACAGGGCACCAGGCCAGA	2391
Db	2553	TTTGCTCTGTCCCGACACCTCTGTGACACACGCGCATACAGGGCACCAGGCCAGA	2582
Oy	2392	GATAGCTTTCTTTTGTGCTCTTTCGGCCTCGGTAGTCAGTTTTTCATAGCCTTACAGT	2451
Db	2583	GATAGCTTTCTTTTGTGCTCTTTCGGCCTCGGTAGTCAGTTTTTCATAGCCTTACAGT	2642
Oy	2452	ATCTGGCTTTGTACTGAGAAATAAAACACATTTTCATATTTGGTTAAAAA	2511
Db	2643	ATCTGGCTTTGTACTGAGAAATAAAACACATTTTCATATAAAAAAAAAAAAAA	2702
Oy	2512	AA 2513	
Db	2703	AA 2704	
RESULT 3			
AB073252			
ID	AB073252	standard; cDNA; 2704 BP.	
XX	AB073252;		
XX			
DT	30-SEP-2002	(first entry)	
XX			
DE	Human MAP kinase phosphatase splice form 1 cDNA sequence SEQ ID NO:7.		
XX			
KW	Human: phosphatase; mitogen activated protein kinase phosphatase;		
KW	MAP kinase; enzyme; chromosome 11; single nucleotide polymorphism;		
XX	SNP; gene; ss.		
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	5'UTR	1..93	
FT		/*tag- a	
FT	CDS	94..1509	
FT		/*tag- b	
FT		/product- "MAP kinase phosphatase splice form 1"	
FT	variation	replace(577,A)	
FT		/*tag- c	
FT	variation	/standard_name=	"single nucleotide polymorphism (SNP)"
FT		replace(1451,G)	
FT		/*tag- d	
FT	3'UTR	/standard_name=	"single nucleotide polymorphism (SNP)"
FT		1510..2704	
FT		/*tag- e	
FT	variation	replace(2641,A)	
FT		/*tag- f	
FT		/standard_name=	"single nucleotide polymorphism (SNP)"
XX			
FN	MO200242436-A2.		
XX			
DD	30-MAY-2002.		
XX			
PE	07-NOV-2001; 2001MO-US42995.		
XX			
XX	20-NOV-2000; 2000US-0715177.		
PR	18-JAN-2001; 2001US-0761640.		
XX			
PA	(PEKE) PE CORP NY.		
XX			
PI	Wei M, Ketchum KA, Di Francesco V, Beasley EM;		
XX			
DR	WPI: 2002-575237/61.		
DR	P-PSDB; ABP51653.		
XX			
PT	Novel isolated human phosphatase peptide useful for treating disorder		
PT	characterized by absence of, inappropriate or unwanted expression of		
PT	the phosphatase protein, and as immunogens to raise antibodies -		
XX			
SS	Claim 1: Fig 3A; 85pp; English.		

[illegible]

Db 787 TACCAGAGAGACTGAACTCCGAACAGAGCTGGCTCATATGATGAGAGGCTATGCGCCGAC 846
QY 689 -----GTTCTCAGAAACAGGAGCAGATG 711
Db 847 CTGGAGTCTTCGGGGCTCCAGCGCCGAGCCTGGCGGGCTCTCAGAAACAGAGAGCAGATG 906
QY 712 GAGCAGGCGATCCGTGCTGAGCTGTGGAAGTGTGGATGTGATGATGACCTCGAGAGCTGTC 771
Db 907 GAGCAGGCGATCCGTGCTGAGCTGTGGAAGTGTGGATGTGATGATGACCTCGAGAGCTGTC 966
QY 772 ACTTCCAAAGAGATCCGCCAGGCTGTGAGAGCTGGGCTGCCCTCCAGAGCTGAC 831
Db 967 ACTTCCAAAGAGATCCGCCAGGCTGTGAGAGCTGGGCTGCCCTCCAGAGCTGAC 1026
QY 832 CGTGAATTCATGCAACAACAGATGCTGTGCTGGGAGAGAGAGGAGGAGAGGAGCTCCCGC 891
Db 1027 CGTGAATTCATGCAACAACAGATGCTGTGCTGGGAGAGAGAGGAGGAGAGGAGCTCCCGC 1086
QY 892 ATCTTCCCGCCACCTTACCTGGGCTCAGAGTGAACGACGAACAACTGGAGAGCTGCAG 951
Db 1087 ATCTTCCCGCCACCTTACCTGGGCTCAGAGTGAACGACGAACAACTGGAGAGCTGCAG 1146
QY 952 AGGAACAGGGTACCCACATCTTTGAACATGCGCCGGAGATTGACAACTTCTAACCTGAG 1011
Db 1147 AGGAACAGGGTACCCACATCTTTGAACATGCGCCGGAGATTGACAACTTCTAACCTGAG 1206
QY 1012 CGCTTACCTACCAACATGCGGCTCTGGATGAGAGATGCGCCAGCTGCTGCGCGAC 1071
Db 1207 CGCTTACCTACCAACATGCGGCTCTGGATGAGAGATGCGCCAGCTGCTGCGCGAC 1266
QY 1072 TGGAGAGAGACCCACCGCTTCAATTAGAGCTGCAAGACACAGGAGACCCACGCTGTGTC 1131
Db 1267 TGGAGAGAGACCCACCGCTTCAATTAGAGCTGCAAGACACAGGAGACCCACGCTGTGTC 1326
QY 1132 CACTCAGAGATGGGGGTGACGCGCTCAGAGGGCCACAGTGTGGGCTATGGCTTGAAGAG 1191
Db 1327 CACTCAGAGATGGGGGTGACGCGCTCAGAGGGCCACAGTGTGGGCTATGGCTTGAAGAG 1386
QY 1132 TACGATGATGAGCTGTGAGAGAGGCTGCGCCACGTCAGAGAGCTCGGCGCCATGCGCCGC 1251
Db 1387 TACGATGATGAGCTGTGAGAGAGGCTGCGCCACGTCAGAGAGCTCGGCGCCATGCGCCGC 1446
QY 1252 CCAACCTCTGGGCTCTGCGGCGAGCTGAGATCTAACCGAGGCTCTTACGCGCCAGAAC 1311
Db 1447 CCAACCTCTGGGCTCTGCGGCGAGCTGAGATCTAACCGAGGCTCTTACGCGCCAGAAC 1506
QY 1312 TGAGGGTGTGGGAGAGAAAGTTGTAGGATGGAAGAGAGAGAGAGAGAGAGAGAGAG 1371
Db 1507 TGAGGGTGTGGGAGAGAAAGTTGTAGGATGGAAGAGAGAGAGAGAGAGAGAGAGAG 1566
QY 1372 AGAGCTGGGGCCACGAGGGGCAAGTAAACCTCGAGGGGCTCATGAGTTCATGATCTT 1431
Db 1567 AGAGCTGGGG--CCACGCGCCAGTAAACCTCGAGGGGCTCATGAGTTCATGATCTT 1624
QY 1432 CTGGAGCCTCTCTTGGAGCTGAGAGACCTCAGTAGACAGAGTGAATGCGCAGAGGCT 1491
Db 1625 CTGGAGCCTCTCTT--GGAGGCTGGAGAGACACTCTAG--AGACAGTGAATGCGCAGAGGCT 1682
QY 1492 TCTTCTCCACGAGTCTTACATGAAGAGCTGTGAGAGCTTCCACAGCTTGAAGGA 1551
Db 1683 TCTTCTCCACGAGTCTTACATGAAGAGCTGTGAGAGCTTCCACAGCTTGAAGGA 1742
QY 1552 CCAAGGAGGCGACAGAGTGAACAGAGGGGCTCAGGCTGCGCTGAAATGCCGAGTCAAG 1611
Db 1743 CCAAGGAGGCGACAGAGTGAACAGAGGGGCTCAGGCTGCGCTGAAATGCCGAGTCAAG 1802
QY 1612 TGGTTACCTTCAAGGAGTGCCTGTGGCAACCGAGAGGCTTCCAGAGAGAGG 1671
Db 1803 TGGTTACCTTCAAGGAGTGCCTGTGGCAACCGAGAGGCTTCCAGAGAGAGG 1862
QY 1672 AGCAGGGGAGGGGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1731

Db 1863 AGCAGGGGAGAGGGGACAGGGGACAGGAGAGCCCTGATTTCTCTACGCCAGGTTCCGGA 1922
QY 1732 AGTGTGAGACAGGCGCAGCGTGCATGACATGGAGAGAGAGAGGCGGCTACGCCCTCA 1791
Db 1923 AGTGTGAGACAGGCGCAGCGTGCATGACATGGAGAGAGAGAGAGGCGGCTACGCCCTCA 1982
QY 1792 CACATCCCAACGCTCCCTGACACTGAAGAGATCCACAACTCTTGGAAACACCTTC 1851
Db 1983 CACATCCCAACGCTCCCTGACACTGAAGAGATCCACAACTCTTGGAAACACCTTC 2042
QY 1852 AGCTGTGGCGGACACATTCCTCTCAGCTTCGGCCCATACCCGTCACATACAGCTTAC 1911
Db 2043 AGCTGTGGCGGACACATTCCTCTCAGCTTCGGCCCATACCCGTCACATACAGCTTAC 2102
QY 1912 CTCCACCCCTGTGACATGAGGCTCAGCTCCACCCCTGTGCTACAGCTTACCTCTCA 1971
Db 2103 CTCCACCCCTGTGACATGAGGCTCAGCTCCACCCCTGTGCTACAGCTTACCTCTCA 2162
QY 1972 CAGCTTAAAGTCCAGAGGCTGATGCTGCTGTCAGAGGAGCTCAAGACTTCTAACCTGGA 2031
Db 2163 CAGCTTAAAGTCCAGAGGCTGATGCTGCTGTCAGAGGAGCTCAAGACTTCTAACCTGGA 2222
QY 2032 TGTGTTAAGAGGACTGAAGATACCTTTGGGGCAACAGACCTTATGTTCTCAACT 2091
Db 2223 TGTGTTAAGAGGACTGAAGATACCTTTGGGGCAACAGACCTTATGTTCTCAACT 2282
QY 2092 CTAGCCTGTGACACTCAGCTGTGGCAGCGAATGAACAGAGCTTCCGTCGCAAAAAGG 2151
Db 2283 CTAGCCTGTGACACTCAGCTGTGGCAGCGAATGAACAGAGCTTCCGTCGCAAAAAGG 2342
QY 2152 TCAGGCTGCCACCCCGGCGGCTCCCTGTCACGCTGCTGCTCTCCAGTTCATTCCTGG 2211
Db 2343 TCAGGCTGCCACCCCGGCGGCTCCCTGTCACGCTGCTGCTCTCCAGTTCATTCCTGG 2402
QY 2212 AACCAAGGAGGCGCAGGCAACCAAGTGGCCCCCAAGAGAGAGAGATCCTCAGGCGCCAGC 2271
Db 2403 AACCAAGGAGGCGCAGGCAACCAAGTGGCCCCCAAGAGAGAGAGATCCTCAGGCGCCAGC 2462
QY 2272 GCGGGAGAGGCTGGAAGGCTGTGGAGATGCTGCTCTCTATCCACTCCACGCGTCAAGTGC 2331
Db 2463 GCGGGAGAGGCTGGAAGGCTGTGGAGATGCTGCTCTCTATCCACTCCACGCGTCAAGTGC 2522
QY 2332 TTTGCTGCTGTCCACAGACGCTGTGACACACAGCGAGATCAGAGGCGACAGGCGCAGA 2391
Db 2523 TTTGCTGCTGTCCACAGACGCTGTGACACACAGCGAGATCAGAGGCGACAGGCGCAGA 2582
QY 2392 GATAGTCTCTTTTGTGCTCTTGTGCGCTGTGGCTAGTCAAGTTTTCATGACCTTACAGT 2451
Db 2583 GATAGTCTCTTTTGTGCTCTTGTGCGCTGTGGCTAGTCAAGTTTTCATGACCTTACAGT 2642
QY 2452 ATCTGCTTTGTAAGTGAATAAACATTTTCATATTTGGTTTAAAAA 2511
Db 2643 ATCTGCTTTGTAAGTGAATAAACATTTTCATATTTAAAAA 2702
QY 2512 AA 2513
Db 2703 AA 2704

RESULT 4
AAD36063
ID AAD36063 standard; cDNA; 2618 BP.
AC AAD36063;
XX
DT 09-AUG-2002 (first entry)
XX
DE Murine dual-specificity phosphatase 15 (Dsp-15) cDNA.
XX
KW Murine: dual-specificity phosphatase 15; DSP15; anti-allergic; cytosolic;
KW immunosuppressive; MAP; mitogen activated protein kinase; cancer; enzyme;
KW signal transduction; cell proliferation; Duchenne muscular dystrophy;
KW cell cycle abnormality; graft-versus-host disease; autoimmune disease;

KW metabolic disease; allergy; screening; gene; ss.
 XX Mus musculus.
 OS
 FH Key Location/Qualifiers
 FT CDS 35..1450
 FT /*tag= a
 FT /product= "Murine DSP-15 protein"
 XX
 XX W0200224740-A2.
 XX
 XX 28-MAR-2002.
 XX
 XX 19-SEP-2001; 2001WO-US29406.
 XX
 XX 19-SEP-2000; 2000US-233833P.
 PR 18-SEP-2001; 2001US-0955732.
 PA (CEPT-) CEPTYR INC.
 XX
 XX Luche RM, Wei B;
 PI
 XX
 XX WPI: 2002-394127/42.
 DR P-PSDB: AAE22733.
 XX
 PT New dual-specificity phosphatase 15 polypeptide and polynucleotides,
 PT useful for treating e.g. Duchenne muscular dystrophy, cancer,
 PT graft-versus-host disease, autoimmune diseases, allergies, metabolic
 PT diseases
 XX
 XX Claim 56; Fig 4; 91pp; English.
 PS
 XX
 CC The invention relates to a new isolated dual-specificity phosphatase 15
 CC (DSP-15) polypeptide which retains the ability to dephosphorylate an
 CC activated MAP (mitogen activated protein) kinase. DSPs are phosphatases
 CC that dephosphorylate both phosphotyrosine and phosphothreonine/serine
 CC residues. DSP-15 polypeptides may be used to identify agents that
 CC modulate DSP-15 activity, where such agents may inhibit or enhance signal
 CC transduction via a MAP-kinase cascade, leading to cell proliferation. DSP
 CC polypeptides, modulating agents, and/or polynucleotides encoding the
 CC polypeptides may be used to modulate DSP-15 activity in a patient, and to
 CC ameliorate disorders such as Duchenne muscular dystrophy, cancer, graft-
 CC versus-host disease, autoimmune diseases, allergies, metabolic diseases,
 CC abnormal cell growth, abnormal cell proliferation and cell cycle
 CC abnormalities. DSP-15 alternate form polypeptides are useful in screening
 CC assays for modulators of enzyme activity and/or substrate binding. The
 CC present sequence is murine DSP-15 cDNA.
 CC
 XX
 XX Sequence 2618 BP; 538 A; 857 C; 769 G; 454 T; 0 other;
 SQ
 Query Match 86.2%; Score 2189.6; DB 24; Length 2618;
 Best Local Similarity 92.5%; Pred. No. 0;
 Matches 2425; Conservative 0; Mismatches 4; Indels 193; Gaps 4;

QY 354 GAAGCAGAGAGCAGAGCAGACCTGACCTCATGGTTACAGCTGTGAGCCGACAGA 413
 |||||||
 Db 301 GAAGCAGAGAGAGCAGAGCAGACCTGACCTCATGGTTACAGCTGTGAGCCGACAGA 360
 QY 414 TGACATCGCGCTGGAGCCAGCTGGAGGACACCCCGGCTCCCGGCTCCGCTACTGCT 473
 |||||||
 Db 361 TGACATCGCGCTGGAGCCAGCTGGAGGACACCCCGGCTCCCGGCTCCGCTACTGCT 420
 QY 474 GGTAGTTTCTACACAGAGAGAGAGAGTGTGAGCCAGATGAGAGGTCTCTCTGGGCGT 533
 |||||||
 Db 421 GGTAGTTTCTACACAGAGAGAGAGAGTGTGAGCCAGATGAGAGGTCTCTCTGGGCGT 480
 QY 534 GGATTTCCCTGACAGACACCTCCCGAGCTGCACCTTGAGGCTGGCTGGCTCCGCTTGAG 593
 |||||||
 Db 481 GGATTTCCCTGACAGACACCTCCCGAGCTGCACCTTGAGGCTGGCTGGCTCCGCTTGAG 540
 QY 594 TGACACCCAGGTGTACTTATGATGAGAGAGGAGGCTTACAGCTGACGTGTGGGCAAG 653
 |||||||
 Db 541 TGACACCCAGGTGTACTTATGATGAGAGAGGAGGCTTACAGCTGACGTGTGGGCAAG 600
 QY 654 CCGGATCTTCAAGCCCATCTCCATCCAGCATG----- 688
 |||||||
 Db 601 CCGGATCTTCAAGCCCATCTCCATCCAGCATGTTGGGCCACACTCCAGTATTGCACCA 660
 QY 689 ----- 688
 |||||||
 Db 661 AGCATGTAGAGCAGCTTAAGCAGCGGCTTGTACCGGTTGGAGTGCCTTACCTGGGC 720
 QY 689 ----- 688
 |||||||
 Db 721 CAGCCACTACAGAGAGAGTGAATCCGACAGAGTGTCTCATATGATGGAGGCTAT 780
 QY 689 ----- 688
 |||||||
 Db 781 GAGCCAGCTGAGTCTGCGGCTCCAGGCGCAGGCTGGCGGCTCTAGAAACAGA 840
 QY 705 GCAGATGAGAGAGGAGGATCCGCTGAGCTGTGAAAGTGTGATGCAGTACCTGGA 764
 |||||||
 Db 841 GCAGATGAGAGAGGAGGATCCGCTGAGCTGTGAAAGTGTGATGCAGTACCTGGA 900
 QY 765 GAGTGTCACTTCCAAAGAGATCCGCGAGCTGTGAGGCTGGGGCTCCCGCTCCA 824
 |||||||
 Db 901 GAGTGTCACTTCCAAAGAGATCCGCGAGCTGTGAGGCTGGGGCTCCCGCTCCA 960
 QY 825 GCAGTACCGTGTACTTATGAGAGAGAGATGCTGTGCTGTGGTGGACAGCGGACCGAC 884
 |||||||
 Db 961 GCAGTACCGTGTACTTATGAGAGAGAGATGCTGTGCTGTGGTGGACAGCGGACCGAC 1020
 QY 885 CTCCGCGATCTTCCCGCAGCTTACCTGGGCTCAGAGTGAAGAGCAAGCAACTGGAGGA 944
 |||||||
 Db 1021 CTCCGCGATCTTCCCGCAGCTTACCTGGGCTCAGAGTGAAGAGCAAGCAACTGGAGGA 1080
 QY 945 GCTGCAGAGAGAGAGGCTCACCATCTTGAACATGGCCCGGAGATGTGAACATTCTA 1004
 |||||||
 Db 1081 GCTGCAGAGAGAGAGGCTCACCATCTTGAACATGGCCCGGAGATGTGAACATTCTA 1140
 QY 1005 CCTGTAGGCTTTCACCTTACCAATATGCTGCTCTGGGATGAGAGTGGCCCACTCT 1064
 |||||||
 Db 1141 CCTGTAGGCTTTCACCTTACCAATATGCTGCTCTGGGATGAGAGTGGCCCACTCT 1200
 QY 1065 GCGGACATGAGAGAGAGAGCGGCTTATGAGGCTCAAGAGAGACAGGGGACACCACT 1124
 |||||||
 Db 1201 GCGGACATGAGAGAGAGAGCGGCTTATGAGGCTCAAGAGAGACAGGGGACACCACT 1260
 QY 1125 GCTGTCCACTGTCAAGATGGGCTGAGCGCTCAGCGGCTCAAGTGTGGCTATGCGCAT 1184
 |||||||
 Db 1261 GCTGTCCACTGTCAAGATGGGCTGAGCGCTCAGCGGCTCAAGTGTGGCTATGCGCAT 1320
 QY 1185 GAAGCAGTACCAATGACGCTGTGAGAGAGGCGCTGCGCCACGCTGCAGAGAGCTTCGGCCAT 1244
 |||||||
 Db 1321 GAAGCAGTACCAATGACGCTGTGAGAGAGGCGCTGCGCCACGCTGCAGAGAGCTTCGGCCAT 1380

QY	1245	CGCGCGCCCAACCCGCGCTCCTCGCGCCAGCTGCAAGATACAGAGGCATCCTGCAGGC	1304
Db	1381	CGCGCCGCCCAACCCCTGGCTCTCTGGCCAGCTGCAGATCTACAGGSCATCTCGACGC	1440
QY	1305	CAGAACCTGAGGGTGTGGGGAGAGAACAGTGTAGGCATGGAAAGAGCCAGCGACGCC	1364
Db	1441	CAGAACCTGAGGGGTGGGGAGAGAACAGTGTAGGCATGGAAAGAGCCAGCGACGCC	1500
QY	1365	CGAAAGAAAGAGCTGGGGGCCACGGGGCAGGTATTAACCTCGAGGGGTCAATAGAGTCAT	1424
Db	1501	CGAAAGAAAGAGCTGGG--CCACGGCCAGGTATTAACCTCGAGGGGTCAATAGAGTCAT	1558
QY	1425	CAGTCTTCTGGAGCCCTCTTGGGAGCTGGAGAGCACTCAAGTACACAGTACATGCCA	1484
Db	1559	CAGTCTTCTGGAGCCCTCTT--GGAGCTGGAGAGCACTCAAG--AGACAGTACATGCCA	1616
QY	1485	GAGTCTTCTCTTCCACAGATTTCAATGAAGAGCTCTGACAGCTTCCACAGCTT	1544
Db	1617	GAGTCTTCTCTTCCACAGATTTCAATGAAGAGCCCTGACAGCTTCCACAGCTT	1676
QY	1545	GCAAGAGCAAGGGAGGGCCAGCAGGTGGAGAGAGGGGCCCTCAGCTCGCCCTAAGTCCGC	1604
Db	1677	GCAAGAGCAAGGGAGGGCCAGCAGGTGGAGAGGGGCCCTCAGCTCTGCCCTGAATCCCGC	1736
QY	1605	CAGTCAAGTGTTFACCTCCAGGGCAGTGCCTGTGTGGCCAAACCGAGCCAGGCTTCAG	1664
Db	1737	CAGTCAAGTGTTFACCTCCAGGGCAGTGCCTGTGTGGCCAAACCGAGCCAGGCTTCAG	1796
QY	1665	GAGGAGAGAGGGGGGCGAGGGGGCGAGGGGAGAGGCCCTGTGATTTCTCTACGCCACAG	1724
Db	1797	GAGGAGAGAGGGGGGCGAGGGGGCGAGGGGAGAGGCCCTGTGATTTCTCTACGCCACAG	1856
QY	1725	TTCCGGAAGGTGTGAGACAGGCCAGCGTGCATGACAGTGAAGAGAGGGCGAGGCTCGA	1784
Db	1857	TTCCGGAAGGTGTGAGACAGGCCAGCGTGCATGACAGTGAAGAGAGGGCGAGGCTCGA	1916
QY	1785	GCCCTCACACATGGCCACAGCTGCCCTGTGACACTGAAGAGATCCACACCTCTGGAGAA	1844
Db	1917	GCCCTCACACATGGCCACAGCTGCCCTGTGACACTGAAGAGATCCACACCTCTGGAGAA	1976
QY	1845	CACCCCTCAGTCTTATCCCGACACATATTCCTTCACACTCGGCCCATACCGTCACTACA	1904
Db	1977	CACCCCTCAGTCTTATCCCGACACATATTCCTTCACACTCGGCCCATACCGTCACTACA	2036
QY	1905	GCGTCACTCCACACCCCTGTCACTAGGGCTCACTCTCCACCCCTGTCACTACAGCTCA	1964
Db	2037	GCGTCACTCCACACCCCTGTCACTAGGGCTCACTCTCCACCCCTGTCACTACAGCTCA	2096
QY	1965	CGTCTTACAGCTTAAGTCCAGAGGCCATGTGCTGTCAAGGGCTCAAGACTTTCTA	2024
Db	2097	CGTCTTACAGCTTAAGTCCAGAGGCCATGTGCTGTGTCAAGGGCTCAAGACTTTCTA	2156
QY	2025	ACTGGAGTGTGTAGAGGGATGAAGGTACTTTGGGGGCAACAGCACTAGTTCAAT	2084
Db	2157	ACTGGAGTGTGTGTAGAGGGATGAAGGTACTTTGGGGGCAACAGCACTAGTTCAAT	2216
QY	2085	GTCAACTGTAGCCCGACACATCACTCACTGTGGCAGGAAATGAAGAAACAGACTCCCGTCA	2144
Db	2217	GTCAACTGTAGCCCTGCACACTCACTGTGGCAGGAAATGAAGAAACAGACTCCCGTCA	2276
QY	2145	AAAAGGTGTACGCTCCACACCCCGCCCTCCTGTCACTCTGTCTCTCCAGTTCA	2204
Db	2277	AAAAGGTGTACGCTCCACACCCCGCCCTCCTGTCACTCTGTCTCTCCAGTTCA	2336
QY	2205	TTTCTGTGAACACGACGAGCCACAGCAACCAATGGCCCCCAAGAGCAGGATCCTCAAG	2264
Db	2337	TTTCTGTGAACACGACGAGCCACAGCAACCAATGGCCCCCAAGAGCAGGATCCTCAAG	2396
QY	2265	CCCAGACCGCGGAGGCTGGAAAGGGTGGAGATCGTTCCTTCATTCACATCCACCGT	2324
Db	2397	CCCAGACCGCGGAGGCTGGAAAGGGTGGAGATCGTTCCTTCATTCACATCCACCGT	2456
QY	2325	CCAGGCTTTGTGCTGTCCCGACAGCTCCTGTGACACAGCCAGATACACAGGCGACCA	2384

Accession	Sequence	Position
Db	CGAGTCTTTCCTGCTGTGCCCCAGACCTCTTGACACACGCGCAGATCATCAGGCGACCA	2518
Qy	GGCCAGAGATAGTCTTCTTTTGTCCCTTTTGCCCTCTGGCAGTAGCATTTTTCATATACC	2444
Db	GCCAGAGATAGTCTTCTTTTGTCCCTTTTGCCCTCTGGCAGTAGCATTTTTCATATACC	2518
Qy	TTACAGTATCTGCGTTTGCTGCTGGAATTAACACATTTTC	2486
Db	TTACAGTATCTGCGTTTGCTGCTGGAATTAACACATTTTC	2518
Qy	TTACAGTATCTGCGTTTGCTGCTGGAATTAACACATTTTC	2486
Db	TTACAGTATCTGCGTTTGCTGCTGGAATTAACACATTTTC	2518

XX	RESULT 5
ID	AB073250
XX	AB073250 standard; cDNA; 2852 BP.
XX	AB073250;
XX	
DT	30-SEP-2002 (first entry)
XX	
DE	Human MAP kinase phosphatase splice form 2 cDNA sequence SEQ ID NO:2.
XX	
KW	Human; phosphatase; mitogen activated protein kinase phosphatase;
XX	MAP kinase; enzyme; chromosome 11; gene; ss.
OS	Homo sapiens.
XX	
EH	Key
FT	Location/Qualifiers
FT	5'UTR
FT	1..56
FT	/*tag= a
FT	57..2036
FT	/*tag= b
FT	/*product= "MAP kinase phosphatase splice form 2"
FT	2037..2852
FT	/*tag= c
XX	
PN	WO200242436-A2.
XX	
PD	30-MAY-2002.
XX	
PF	07-NOV-2001; 2001WO-US42995.
XX	
PR	20-NOV-2000; 2000US-0715177.
PR	18-JAN-2001; 2001US-0761640.
XX	
PA	(PEKE) PE CORP NY.
XX	
PI	Wei M, Ketchum KA, Di Francesco V, Beasley EM;
XX	
DR	WPI: 2002-575237/61.
XX	
DR	P-PSDB; ABP51654.
XX	
PT	Novel isolated human phosphatase peptide useful for treating disorder
PT	characterized by absence of, inappropriate or unwanted expression of
PT	the phosphatase protein, and as immunogens to raise antibodies
XX	
PS	Claim 1; Fig 1B; 85pp; English.
XX	
CC	The present invention describes an isolated human phosphatase peptide
CC	(I). (I) can be used for identifying a modulator of (I) by contacting
CC	(I) with an agent and determining if the agent has modulated the
CC	function or activity of (I). (I) is useful for identifying an agent that
CC	binds to (I), by contacting (I) with an agent and assaying the contacted
CC	mixture to determine whether a complex is formed with the agent bound
CC	(I). The human phosphatases from the present invention are mitogen
CC	activated protein (MAP) kinase phosphatases. These human MAP kinase
CC	phosphatases are located on chromosome 11. (I) and the polynucleotide
CC	sequences encoding (I) can be used in gene therapy. The present sequences
CC	encodes human MAP kinase phosphatase splice form 2 from the present
CC	invention.
XX	
SQ	Sequence 2852 BP; 674 A; 895 C; 807 G; 476 T; 0 other;

Db 2098 ACCCTCAGCTGTGTGTCGCGACACATTTCCCTCAGCTCCGCCCATACCCCTCAGCTACAG 2157
 QY 1906 CCTCAGCTCCACCCCTGTACAGGCTCAGCTCCACCCCTGTACAGCTACAGCTCAG 1965
 Db 2158 CCTCAGCTCCACCCCTGTACAGGCTCAGCTCCACCCCTGTACAGCTACAGCTCAG 2217
 QY 1966 CTCTCAGCTCCTTAAGTCCAGGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 2025
 Db 2218 CTCTCAGCTCCTTAAGTCCAGGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 2277
 QY 2026 CTGGATGTGTAGAGGAGCTAAGTACCTTTGGGGCAACAGCAGCTAGTTTCATTC 2085
 Db 2278 CTGGATGTGTAGAGGAGCTAAGTACCTTTGGGGCAACAGCAGCTAGTTTCATTC 2337
 QY 2086 TCACTCTAGCCCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 2145
 Db 2338 TCACTCTAGCCCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 2397
 QY 2146 AAAGGTCAGGCTCCACCCCTCCCTCCCTCAGCTCAGCTCAGCTCAGCTCAG 2205
 Db 2398 AAAGGTCAGGCTCCACCCCTCCCTCCCTCAGCTCAGCTCAGCTCAGCTCAG 2457
 QY 2206 TCCTGGAACAGCAGGAGCAACAGTGTGAGCCCAAGAGGAGAGATCTCAGGC 2265
 Db 2458 TCCTGGAACAGCAGGAGCAACAGTGTGAGCCCAAGAGGAGATCTCAGGC 2517
 QY 2266 CCCAGCCGCGGAGGAGGAGGAGGAGTGGAGTCCCTCAGCTCAGCTCAGCTCAG 2325
 Db 2518 CCCAGCCGCGGAGGAGGAGGAGGAGTGGAGTCCCTCAGCTCAGCTCAGCTCAG 2577
 QY 2326 CAGCTCTTCTGCTGTCCTCCAGACCTCTGTACACACAGGAGTTCACAGGACACAG 2385
 Db 2578 CAGCTCTTCTGCTGTCCTCCAGACCTCTGTACACACAGGAGTTCACAGGACACAG 2637
 QY 2386 GCCAGATATGTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 2445
 Db 2638 GCCAGATATGTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 2697
 QY 2446 TACAGTATCTGCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 2505
 Db 2698 TACAGTATCTGCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 2757
 QY 2506 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2540
 Db 2758 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2792

RESULT 6
 AAF30485
 ID AAF30485 standard; cDNA; 2781 BP.
 AC AAF30485;
 XX
 DT 29-MAY-2001 (first entry)
 XX
 DE Human protein phosphatase and kinase protein-10 cDNA 5039718CB1.
 XX
 KW Protein phosphatase and kinase protein; PPHK-10; human;
 KW gastrointestinal disorder; immune system disorder;
 KW neurological disorder; cell proliferative disorder; cancer;
 KW diagnosis; therapy; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FT CDS Location/Qualifiers
 FT 87..2066
 FT sig_peptide /*lag= a
 FT 87..143
 FT mat_peptide /*lag= b
 FT 144..2063
 FT /*tag= c
 FT 219..279
 FT misc_feature /*tag= d

FT FT /note= "unique fragment"
 FT misc_feature 921..980
 FT /*lag= e
 FT /*note= "unique fragment"
 PN WO200120004-A2.
 XX
 XX 22-MAR-2001.
 PD
 PF 14-SEP-2000; 2000MO-US25515.
 PR 15-SEP-1999; 99US-0154141.
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 XX Yue H, Tang YT, Bandman O, Hillman JL, Baughn MR, Azimzal Y;
 PI Lu DAM;
 PI
 DR WPI; 2001-244811/25.
 DR P-PSDB; AAB20331.
 XX
 PS Claim 5; Page 101-102; 103pp; English.
 CC The present sequence is that of cDNA encoding novel human
 CC protein phosphatase and kinase protein PPHK-10 (see AAB20331).
 CC The cDNA was initially identified in Incyte Clone ID No. 5039718CB1,
 CC (as a fraction of total tissues expressing PPHK-10
 CC reproductive (0.343), gastrointestinal (0.194) and haematopoietic
 CC or immune (0.134). Diseases or conditions associated with tissues
 CC expressing PPHK-10 (as a fraction of total tissues expressing
 CC cell proliferation (0.090). The encoded protein shows homology to
 CC Drosophila melanogaster MAP kinase phosphatase. The invention
 CC provides human PPHK-1 to -11 polypeptides (see AAB20332-32) and
 CC polynucleotides (see AAF30476-86). It also provides expression
 CC vectors, host cells, antibodies, agonists and antagonists, as well
 CC as methods for diagnosing, treating or preventing disorders
 CC associated with expression of PPHK, including gastrointestinal
 CC disorders, immune system disorders, neurological disorders and cell
 CC proliferative disorders, including cancer.
 SQ Sequence 2781 BP; 576 A; 906 C; 820 G; 479 T; 0 other;
 Query Match 83.3%; Score 2115; DB 22; Length 2781;
 Best Local Similarity 89.3%; Pred. No. 0;
 Matches 2470; Conservative 0; Mismatches 5; Indels 292; Gaps 6;
 QY 10 GCGGTCACAGACTGTCGCGGAGGTTGAGGAGGCGCTGCTCCAGCCAGG 69
 Db 10 GCGGTCACAGACTGTCGCGGAGGTTGAGGAGGCGCTGCTCCAGCCAGG 68
 QY 70 TGTCTCGGCTGCTGCTCAATGAGCTGTCAGTGAAGCGTTGCGCCCGGAGCGGCG 129
 Db 69 TGTCTCGGCTGCTGCTCAATGAGCTGTCAGTGAAGCGTTGCGCCCGGAGCGGCG 128
 QY 130 GCTTCAGGCGGCTGCTGCTGAGCAAGCGGCTGTCAGCAAGAGTGCATCCAGCA 189
 Db 129 GCTTCAGGCGGCTGCTGCTGAGCAAGCGGCTGTCAGCAAGAGTGCATCCAGCA 188
 QY 190 AGGCAAGCTTTGCGGCTGCTGCTGAGCAAGCGGCTGTCAGCAAGAGTGCATCCAGCA 249
 Db 189 AGGCAAGCTTTGCGGCTGCTGCTGAGCAAGCGGCTGTCAGCAAGAGTGCATCCAGCA 248
 QY 250 GATGATGAGCAGAGGAGGAGTTCAGGCAAGAGGCGGCTGTCAGCAAGAGTGCATCCAGCA 309
 Db 249 GATGATGAGCAGAGGAGGAGTTCAGGCAAGAGGCGGCTGTCAGCAAGAGTGCATCCAGCA 308
 QY 310 CACGGGACAGACAGACTTGGGCAAGAGTCCAGAGTCCCAAGAAAGCAGAGAGCAG 369

|||||
Db 309 CACGGGACACAGACTGCGGCAAGATCCAGAGTCCCAAGAACAGAGAGAGCAG 368
Qy 370 AGGAGACACCTGCACCTATGATAGCTGCTAGAGCCGCGAGATATCCGCTGGCA 429
Db 369 AGGAGACACCTGCACCTATGATAGCTGCTAGAGCCGCGAGATATCCGCTGGCA 428
Qy 430 GCCAGCTGAGAGGACCCCGGCTCCCGGCTCCGCTACCTAGCTGGTATTCTACAGA 489
Db 429 GCCAGCTGAGAGGACCCCGGCTCCCGGCTCCGCTACCTAGCTGGTATTCTACAGA 488
Qy 490 GAAAGAGAGGCTTGAGGACAGATGAGAGGCTCTCTGGGCGTGATTTCCCTGACAGC 549
Db 489 GAAAGAGAGGCTTGAGGACAGATGAGAGGCTCTCTGGGCGTGATTTCCCTGACAGC 548
Qy 550 AGCTCCCCCAGCTGACACCTGGGCTGCTTTGGCCCTCTGGAGTGAACACCAGGTGTAC 609
Db 549 AGCTCCCCCAGCTGACACCTGGGCTGCTTTGGCCCTCTGGAGTGAACACCAGGTGTAC 608
Qy 610 TTAGATGAGAGCGGGGGCTTGAGCTGACGTCGTGGTGGGCAAGCCGGATCTTCAAGGCC 669
Db 609 TTAGATGAGAGCGGGGGCTTGAGCTGACGTCGTGGTGGGCAAGCCGGATCTTCAAGGCC 668
Qy 670 ATCTCCATCCAGACCATGT----- 688
Db 669 ATCTCCATCCAGACCATGTGGGCGCACACTCCAGTATTGACACCAAGCATGTGAGGCAGCT 728
Qy 689 ----- 688
Db 729 CTAGGACAGCGGCTGTGTACCGGGTGAGTGCCCTCACCTGGGCGACCATCCAGGAG 788
Qy 689 ----- 688
Db 789 AGACTGAACCTCCGAACAGAGCTGCCTCAATGAGTGAAGGCTATGGCCGACCTGAGTCT 848
Qy 689 -----GTTCTTCAAGAACAGAGACAGATGAGAGCG 720
Db 849 CTGCGGCTCCAGTGCAGCGAGCTGGCGGTCTCAGAAACAGAGACAGATGAGAGCG 908
Qy 721 ATCCCTGCTGAGCTGTGGAAGGTGTTGATGTCACTGACCTGTGAGAGTGTCACTTCCAAA 780
Db 909 ATCCCTGCTGAGCTGTGGAAGGTGTTGATGTCACTGACCTGTGAGAGTGTCACTTCCAAA 968
Qy 781 GAGATCCGCGAGGCTCTGAGAGCTGGGCTGGGCTCCCTCCAGCAGTACCTGACTTTC 840
Db 969 GAGATCCGCGAGGCTCTGAGAGCTGGGCTGGGCTCCCTCCAGCAGTACCTGACTTTC 1028
Qy 841 ATGCACAACCAAGATGCTGCTGTGTGGCAGACGCGGACCGAGCTCCCGCATTTTCCC 900
Db 1029 ATGCACAACCAAGATGCTGCTGTGTGGCAGACGCGGACCGAGCTCCCGCATTTTCCC 1088
Qy 901 CACCTCTACCTGGGCTGAGAGTGAAGCAGCAACCTGAGAGAGCTGACAGAGAACAG 960
Db 1089 CACCTCTACCTGGGCTGAGAGTGAAGCAGCAACCTGAGAGAGCTGACAGAGAACAG 1148
Qy 961 GTACCCACATCTTGAATATGGCCCGGAGATGACAACTTCTACCTGAGAGGCTTCAAC 1020
Db 1149 GTACCCACATCTTGAATATGGCCCGGAGATTGACAACTTCTACCTGAGAGGCTTCAAC 1208
Qy 1021 TACCACAATGTGCGGCTGTGGGATGAGAGTGGCCAGCTGCTCCGCACTGGAAGAG 1080
Db 1209 TACCACAATGTGCGGCTGTGGGATGAGAGTGGCCAGCTGCTCCGCACTGGAAGAG 1268
Qy 1081 AGCAGCCGCTTATGAGGCTGCAAGAGCAGAGGCAACCAGCTGCTGTGCTCACTGCAAG 1140
Db 1269 AGCAGCCGCTTATGAGGCTGCAAGAGCAGAGGCAACCAGCTGCTGTGCTCACTGCAAG 1328
Qy 1141 ATGGGCTGAGCGGCTGACGCGCCACAGTGTGGCTTATGACATGAGCAGTACAGAAAGC 1200
Db 1329 ATGGGCTGAGCGGCTGACGCGCCACAGTGTGGCTTATGACATGAGCAGTACAGAAAGC 1388
Qy 1201 AGCCTGAGAGCGGCTGCGCCACGTGACAGAGCTCGGCGCCATGCGCGCCCAACCT 1260
|||||

Db 1389 AGCCTGAGAGAGGCGCTGCGCCAGCTGACAGAGCTCCGCGCCATGCGCGCCCAACCT 1448
Qy 1261 GGGTCTCTGGCGCAGCTGACATGTACAGAGGCAATCCGAGC----- 1302
Db 1449 GGGTCTCTGGCGCAGCTGACATGTACAGAGGCAATCCGAGGCGCAGCGCCAGAGCAT 1508
Qy 1303 ----- 1302
Db 1509 GTCGTGGAGAGAAAGTGGGTGGGTCTCCCAAGAGAGCAACCCAGCCCTGAAGTCTCT 1568
Qy 1303 -----GCCAGACCTGAGGCTGTGGGAGAGAGAGTTGTAGCC 1342
Db 1569 ACACCATTCACACCTCTTCCCGCAAGAACCTGAGGGTGGTGGGAGAGAAAGTTGTAGCC 1628
Qy 1343 ATGGAAGAGAGCGAGGCGAGCCCGCAAGAGACCTGGGGCCACCGGGCAGCTATAAAC 1402
Db 1629 ATGGAAGAGAGCGAGGCGAGCCCGCAAGAGAGCCTGGG--CCACGGCCAGCTATAAAC 1686
Qy 1403 TCCGAGGGGTATGAGGTCATCATAGTCTTGTGAGCCCTCTGGAGCTGAGAGCACC 1462
Db 1687 TCCGAGGGGTATGAGGTCATCATAGTCTTGTGAGCCCTCTT--GGAGCTGAGAGCACC 1745
Qy 1463 TCAGTAGACAGTGAATGACATGACAGAGTCTTCTCCAGAGTCTTCAATGAAGAGCC 1522
Db 1746 TCAG--AGACAGTGAATGACATGACAGAGTCTTCTTCCAGAGTCTTCAATGAAGAGCC 1804
Qy 1523 TCTGAGCGCTTCCACAGCTTGGCAAGAGCAAGGAGGAGCCAGCAGTGTGACAGGGGCC 1582
Db 1805 TCTGAGCGCTTCCACAGCTTGGCAAGAGCAAGGAGGAGCCAGCAGTGTGACAGGGGCC 1864
Qy 1583 TCAGCTGCTCTGAAGTCCCGCAGTCACTGATGTTACCTTCAGGCGAGTGCCTGTGTGGC 1642
Db 1865 TCAGCTGCTCTGAAGTCCCGCAGTCACTGATGTTACCTTCAGGCGAGTGCCTGTGTGGC 1924
Qy 1643 CAACCGAGCCAGGCTTCCAGAGACAGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGCC 1702
Db 1925 CAACCGAGCCAGGCTTCCAGAGACAGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGCC 1984
Qy 1703 CTGCACTTCTCTACGGCAGGTTCCGGAAGGTGTGAGACAGGCTGACGTCATGAGAC 1762
Db 1985 CTGCACTTCTCTACGGCAGGTTCCGGAAGGTGTGAGACAGGCTGACGTCATGAGAC 2044
Qy 1763 TGGAGAGAGGAGGCGAGGCTGACAGCTTACACATGCCCAGCCTCCCTGACACTGAAGAG 1822
Db 2045 TGGAGAGAGGAGGCGAGGCTGACAGCTTACACATGCCCAGCCTCCCTGACACTGAAGAG 2104
Qy 1823 GATCCACAACCTCTTGAAGAACACCTGACGTCTGTGGCCGACACATTCCTCTCAGCT 1882
Db 2105 GATCCACAACCTCTTGAAGAACACCTGACGTCTGTGGCCGACACATTCCTCTCAGCT 2164
Qy 1883 CGGCGCCATACCGCTACATGACGCTCAGCTCCAGCCCTGTCATGAGGCTCAGCTCC 1942
Db 2165 CGGCGCCATACCGCTACATGACGCTCAGCTCCAGCCCTGTCATGAGGCTCAGCTCC 2224
Qy 1943 CACCCCTGTCACTAGAGCTCACTCACTACAGCTTAACTGCCAGGCGCACTGTGCTG 2002
Db 2225 CACCCCTGTCACTAGAGCTCACTCACTACAGCTTAACTGCCAGGCGCACTGTGCTG 2284
Qy 2003 TCCAGGSGCTCAAGACTTTCCTAAGTGGATGTGTAAGAGGAGTGAAGTACCTTTGGGG 2062
Db 2285 TCCAGGSGCTCAAGACTTTCCTAAGTGGATGTGTAAGAGGAGTGAAGTACCTTTGGGG 2344
Qy 2063 GCAACAGACCCCTAATTCATTTCACTGAGCCCTGACACATCCTGATGAGCAGGAA 2122
Db 2345 GCAACAGACCCCTAATTCATTTCACTGAGCCCTGACACATCCTGATGAGCAGGAA 2404
Qy 2123 TGAAGACAGAGCTTCCCTGCAAAAAGGCTCAGGCTCCACCCCGGCGCTCCCTGCA 2182
Db 2405 TGAAGACAGAGCTTCCCTGCAAAAAGGCTCAGGCTCCACCCCGGCGCTCCCTGCA 2464
Qy 2183 CCTCTGTCTCTCTCCAGTTCATTCCTGTGAACACAGCCAGGCGCAGGCAACAGTGGCCCC 2242
Db 2465 CCTCTGTCTCTCTCCAGTTCATTCCTGTGAACACAGCCAGGCGCAGGCAACAGTGGCCCC 2524
|||||

QY 2243 AAAGCAGGAGATCTCTAGGCCAGCGGGAGAGCTGGAAGGCTGGCAGATCGCT 2302
 Db 2525 AAAGGAGGAGATCTCTAGGCCAGCGGGAGAGCTGGAAGGCTGGCAGATCGCT 2584
 QY 2303 TCCCTATCCACTCCACCGGCTCCAGGCTCTTGGCTCTCTCCCAACCTCTGTGACAC 2362
 Db 2585 TCCCTATCCACTCCACCGGCTCCAGGCTCTTGGCTCTCTCCCAACCTCTGTGACAC 2644
 QY 2363 CACCCAGATACAGGAGGAGCAGGAGATAGTCTTTTGTCTTGTGCGCTCT 2422
 Db 2645 CACCCAGATACAGGAGGAGCAGGAGATAGTCTTTTGTCTTGTGCGCTCT 2704
 QY 2423 GGCTATCAGTTTTCATATAGCCTTACAGTATCTGCTTGTACTAGAAATTAACACAT 2482
 Db 2705 GGCTATCAGTTTTCATATAGCCTTACAGTATCTGCTTGTACTAGAAATTAACACAT 2764
 QY 2483 TTTCATA 2489
 Db 2765 TTTCATA 2771

RESULT 7

AAD36061
 ID AAD36061 standard; cDNA; 2718 BP.

AC AAD36061;

DT 09-AUG-2002 (first entry)

DE Human dual-specificity phosphatase 15 (DSP-15) cDNA.

KW Human; dual-specificity phosphatase 15; DSP-15; anti-allergic; cytosolic;
 immunosuppressive; MAP; mitogen activated protein kinase; cancer; enzyme;
 signal transduction; cell proliferation; Duchenne muscular dystrophy;
 cell cycle abnormality; graft-versus-host disease; autoimmune disease;
 metabolic disease; allergy; screening; chromosome 11q; gene; ss.
 OS Homo sapiens.

FX Key location/Qualifiers
 FT CDS 35..2014
 FT /tag= a
 FT /product= "Human DSP-15 protein"

PN WO200224740-A2.

PD 28-MAR-2002.

PF 19-SEP-2001; 2001WO-US29406.

PR 19-SEP-2000; 2000US-233833P.

PR 18-SEP-2001; 2001US-0955732.

PA (CEPT-) CEPTYR INC.

PI Luche RM, Wei B;

DR WPL; 2002-394127/42.

DR P-PSDB; AAE22729.

XX New dual-specificity phosphatase 15 polypeptide and polynucleotides,
 useful for treating e.g. Duchenne muscular dystrophy, cancer,
 diseases - host disease, autoimmune diseases, allergies, metabolic
 diseases

XX Claim 7; Fig 1; 91pp; English.

CC The invention relates to a new isolated dual-specificity phosphatase 15
 CC (DSP-15) polypeptide which retains the ability to dephosphorylate an
 CC activated MAP (mitogen activated protein) kinase. DSPs are phosphatases
 CC that dephosphorylate both phosphotyrosine and phosphothreonine/serine
 CC residues. DSP-15 polypeptides may be used to identify agents that

CC modulate DSP-15 activity, where such agents may inhibit or enhance signal
 CC transduction via a MAP-kinase cascade, leading to cell proliferation. DSP
 CC polypeptides modulating agents, and/or polynucleotides encoding the
 CC polypeptides may be used to modulate DSP-15 activity in a patient, and to
 CC ameliorate disorders such as Duchenne muscular dystrophy, cancer, graft-
 CC versus-host disease, autoimmune diseases, allergies, metabolic diseases,
 CC abnormal cell growth, abnormal cell proliferation and cell cycle
 CC abnormalities. DSP-15 alternate form polypeptides are useful in screening
 CC assays for modulators of enzyme activity and/or substrate binding. The
 CC present sequence is human DSP-15 cDNA. Human DSP-15 gene is located on
 CC chromosome 11q.

SO Sequence 2718 BP; 559 A; 893 C; 795 G; 471 T; 0 other;

Query Match 82.0%; Score 2083.6; DB 24; Length 2718;
 Best Local Similarity 89.2%; Pred. No. 0;
 Matches 2427; Conservative 0; Mismatches 4; Indels 291; Gaps 5;

QY 54 CCGGTGCCAGCCAGGTCCTGCGGCTGCTCCATGAGCCCTGTCACAGTAGGCGCTTC 113
 Db 1 CCGGTGCCAGCCAGGTCCTGCGGCTGCTCCATGAGCCCTGTCACAGTAGGCGCTTC 60
 QY 114 GCCCGCGGAGGCGGCTCCAGCGCCCTGGGCGCTGGACACAGGCGCTCCAGCGAAG 173
 Db 61 GCCCGCGGAGGCGGCTCCAGCGCCCTGGGCGCTGGACACAGGCGCTCCAGCGAAG 120
 QY 174 GAGTCGACTCCAGCAGAGGACAGGCTTTGCGTGTCTGCGTGGGCTGTCGAGACTGCA 233
 Db 121 GAGTCGACTCCAGCAGAGGACAGGCTTTGCGTGTCTGCGTGGGCTGTCGAGACTGCA 180
 QY 234 GGATGAGGAGGACATGATGATGACAGCAGAGGCGCATGTTCTGACCCACAGAGAGGCCCC 293
 Db 181 GGATGAGGAGGACATGATGATGACAGCAGAGGCGCATGTTCTGACCCACAGAGAGGCCCC 240
 QY 294 GAGTGAAGGAGGCTCCAGCGGAGCAGACAGACTTCCGGCAGAGATCCAGATCCCA 353
 Db 241 GAGTGAAGGAGGCTCCAGCGGAGCAGACAGACTTCCGGCAGAGATCCAGATCCCA 300
 QY 354 GAAGCAGAGGAGGACAGCAGCAGCTTACAGTGTGAGAGCCGCGCAGGA 413
 Db 301 GAAGCAGAGGAGGACAGCAGCAGCTTACAGTGTGAGAGCCGCGCAGGA 360
 QY 414 TGACATCGGCTGGAGCGCCAGCTGGAAGCACCCGGGCTCCCGGCTCCGCTACCTGCT 473
 Db 361 TGACATCGGCTGGAGCGCCAGCTGGAAGCACCCGGGCTCCCGGCTCCGCTACCTGCT 420
 QY 474 GGTAGTTTCTACAGAGAGAGAGAGGCTTGAGCCAGATGAGACGCTCTCTGGGCGT 533
 Db 421 GGTAGTTTCTACAGAGAGAGAGAGGCTTGAGCCAGATGAGACGCTCTCTGGGCGT 480
 QY 534 GGATTTCCCTGACAGAGCTCCCGCAGCTGACCTGAGGCGCTGCTGCTGAGAG 593
 Db 481 GGATTTCCCTGACAGAGCTCCCGCAGCTGACCTGAGGCGCTGCTGCTGAGAG 540
 QY 594 TGACACCCAGGTTACTTAAATGAGAGAGGCGGCTTACAGCTGAGCTGAGTGGGCAAG 653
 Db 541 TGACACCCAGGTTACTTAAATGAGAGAGGCGGCTTACAGCTGAGCTGAGTGGGCAAG 600
 QY 654 CCGATCTTCAAGCCCATTCCTCCATCCAGACATGT----- 688
 Db 601 CCGATCTTCAAGCCCATTCCTCCATCCAGACATGTGGGCAACATTCACCA 660
 QY 689 ----- 688
 Db 661 AGCATGTAGAGCAGCTTAGGAGCGGCTTGTACCGGCTGGCAGTGCCTTACTTGAGC 720
 QY 689 ----- 688
 Db 721 CACCCACTTACAGAGAGAGACTGACCTCCAGACAGACTGCTCAATGATGAGAGCGTAT 780
 QY 689 -----GTTCTTCAAGACAGGA 704
 Db 781 GGCCGACTGAGGTCTCTGCGGCTCCAGCGCCGAGCCTGGCGGCTCTCGAGAACAGGA 840

Oy 705 GCAGATGAGAGAGCCGATCCGTCGAGCTGAGTGGAAAGTGTGATGTCAGTGACCTGGA 764
|||
Db 841 GCAGATGAGAGAGCCGATCCGTCGAGCTGTGGAAAGTGTGATGTCAGTGACCTGGA 900
Oy 765 GAGTGTCACTTCCAAAAGATCCGACAGGCTCTGGAGCTGCGCCCTGGGGCTCCCTCCCA 824
|||
Db 901 GAGTGTCACTTCCAAAAGATCCGACAGGCTCTGGAGCTGCGCCCTGGGGCTCCCTCCCA 960
Oy 825 GCAGTACCGTGAATTCATCGACACCAATGTGCTGCTGGTGCGACACAGCGGACCGAGC 884
|||
Db 961 GCAGTACCGTGAATTCATCGACACCAATGTGCTGCTGGTGCGACACAGCGGACCGAGC 1020
Oy 885 CTCCCGCATCTTCCCGACCTCTACCTGGGGCTCAGAGTGGAAAGCAGCAAACTGGAGGA 944
|||
Db 1021 CTCCCGCATCTTCCCGACCTCTACCTGGGGCTCAGAGTGGAAAGCAGCAAACTGGAGGA 1080
Oy 945 GCTGCAGAGAGAAACAGAGTCCACCACATCTTGAAACATGGCCCGGAGATGGACAACTTCTA 1004
|||
Db 1081 GCTGCAGAGAGAAACAGAGTCCACCACATCTTGAAACATGGCCCGGAGATGGACAACTTCTA 1140
Oy 1005 CCCTAGCGCTTCACTTACACCAATGTGCGCTCTGGATGAGAGTGGCCACGCTGCT 1064
|||
Db 1141 CCCTAGCGCTTCACTTACACCAATGTGCGCTCTGGATGAGAGTGGCCACGCTGCT 1200
Oy 1065 GCGGCACTGGAAGAGAGAGCAACCGCTTCACTTGGAGCTGCAAGAGCAGAGGACCCACAGT 1124
|||
Db 1201 GCGGCACTGGAAGAGAGAGCAACCGCTTCACTTGGAGCTGCAAGAGCAGAGGACCCACAGT 1260
Oy 1125 GCTGTCCACTGCAAGATGGGCGTCAGCCGCTCAGCGGCAACAGTGTGCGCTATGCCAT 1184
|||
Db 1261 GCTGTCCACTGCAAGATGGGCGTCAGCCGCTCAGCGGCAACAGTGTGCGCTATGCCAT 1320
Oy 1185 GAAGAGTACCAATGACCTGAGAGAGAGCCCTGGCCACGCTGCGAGAGCTCCGCCCAT 1244
|||
Db 1321 GAAGAGTACCAATGACCTGAGAGAGAGCCCTGGCCACGCTGCGAGAGCTCCGCCCAT 1380
Oy 1245 GCGCGGCCCCAACCTGTGCTCTCTGCGCAGCTGAGATCTTACAGGACATCTTACG- 1302
|||
Db 1381 GCGCGGCCCCAACCTGTGCTCTCTGCGCAGCTGAGATCTTACAGGACATCTTACG- 1440
Oy 1303 ----- 1302
Db 1441 CAGCGCCAGAGCCATGTCCTGGAGACAAAGTGGGTGGGCTCTCCAGAGAGACACC 1500
Oy 1303 -----GCCAAGCTGAGGGTGGGGGA 1326
Db 1501 AGCCCTGAAGTCTCTACACCATTCACACTCTTCGCGCAGAACTGAGGGTGGGGGA 1560
Oy 1327 GGAGAAGGTTTATGAGCATGGAAGAGAGCAGGACGCCGAAAGAGAGCTGGGGCCAC 1386
|||
Db 1561 GGAGAAGGTTTATGAGCATGGAAGAGAGCAGGACGCCGAAAGAGAGCTGGGG-CCA 1618
Oy 1387 GGGGACGATTAACCTCCGAGGGGTCTAGAGTCCATCACTCTTCTGAGAGCCCTCTTG 1446
|||
Db 1619 GGGGACGATTAACCTCCGAGGGGTCTAGAGTCCATCACTCTTCTGAGAGCCCTCTT- 1677
Oy 1447 GGAGCTGGAGAGACACTGATGAGACATGACATGACAGAGTCTTCTCTCCACGAGT 1506
|||
Db 1678 GGAGCTGGAGAGACACTGAG-AGACAGTGAATGACAGAGGTCCTCTCTCCACGAGT 1736
Oy 1507 CTTCACATGAAGAGCCTCTGACAGCCCTTCCACAGCTTTCGAAGAGACCAAGGAGGCGAGC 1566
|||
Db 1737 CTTCACATGAAGAGCCTCTGACAGCCCTTCCACAGCTTTCGAAGAGACCAAGGAGGCGAGC 1796
Oy 1567 AGGTGACAGGGGGCTCAGGCTGCGCTGAAGTCCCGGACAGTCACTGCTTACCTCCAGG 1626
|||
Db 1797 AGGTGACAGGGGGCTCAGGCTGCGCTGAAGTCCCGGACAGTCACTGCTTACCTCCAGG 1856
Oy 1627 GCAGTGCCTGTGTGGCCAAACCGGAGAGGCTTCCAGAGAGAGAGAGAGGGGAGGGG 1686
|||
Db 1857 GCAGTGCCTGTGTGGCCAAACCGGAGAGGCTTCCAGAGAGAGAGAGAGGGGAGGGG 1916

Oy 1687 AGGSCAGGAGAGCCCTGCATTTCTCTACGCCCCAGGTTCCGAAAGTGGAGACAGG 1746
|||
Db 1917 AGGSCAGGAGAGCCCTGCATTTCTCTACGCCCCAGGTTCCGAAAGTGGAGACAGG 1976
Oy 1747 CCAGGTCATGACAGTGGAGAGAGGGGAGGAGGCTTGAGCCCTCCACATGCCACGCTC 1806
|||
Db 1977 CCAGGTCATGACAGTGGAGAGAGGGGAGGAGGCTTGAGCCCTCCACATGCCACGCTC 2036
Oy 1807 CCCTACACTGAAGAGATCCCAACACTCTTGGAGAAACACCCCTACGTCGTGTCCGCA 1866
|||
Db 2037 CCCTACACTGAAGAGATCCCAACACTCTTGGAGAAACACCCCTACGTCGTGTCCGCA 2096
Oy 1867 CACATTTCTTCAGGCTCGCCCATACCCGTCACATGAGCCGTCACCTCCACCCCTGTCA 1926
|||
Db 2097 CACATTTCTTCAGGCTCGCCCATACCCGTCACATGAGCCGTCACCTCCACCCCTGTCA 2156
Oy 1927 CTAGGCGCTCACTCCACCCCTGCTACCTACAGGCTTCACTACAGCCCTTAACTGCA 1986
|||
Db 2157 CTAGGCGCTCACTCCACCCCTGCTACCTACAGGCTTCACTACAGCCCTTAACTGCA 2216
Oy 1987 GGGCCATGTCTGCTGTCCAAAGGCTCAAGACTTCTTAACGTGGATGTGTAGAGGACT 2046
|||
Db 2217 GGGCCATGTCTGCTGTCCAAAGGCTCAAGACTTCTTAACGTGGATGTGTAGAGGACT 2276
Oy 2047 GAAGTACCTTTGGGGGCAACAGCACCCCTAGTTTCATTCGTAACCTCTAGCCCTGCACACT 2106
|||
Db 2277 GAAGTACCTTTGGGGGCAACAGCACCCCTAGTTTCATTCGTAACCTCTAGCCCTGCACACT 2336
Oy 2107 CACCTGTGGCAGGAATGAAAAACAGAGCTTCCCGTGCAAAAAAGGCTCAGCCCTCCACCC 2166
|||
Db 2337 CACCTGTGGCAGGAATGAAAAACAGAGCTTCCCGTGCAAAAAAGGCTCAGCCCTCCACCC 2396
Oy 2167 CCGCCCTCTCTGTGACCTCTCTGTCTCTCTCCAGTTTCATTCCTGGAACAGCAGGCGCAG 2226
|||
Db 2397 CCGCCCTCTCTGTGACCTCTCTGTCTCTCTCCAGTTTCATTCCTGGAACAGCAGGCGCAG 2456
Oy 2227 GCAACAGATGGGCCCCAAAGAGCAGAGAGATCTCAGGCCCCAGCGGGAGGCTGGAA 2286
|||
Db 2457 GCAACAGATGGGCCCCAAAGAGCAGAGAGATCTCAGGCCCCAGCGGGAGGCTGGAA 2516
Oy 2287 GGGCTGCGAGATGCTTCTCTATCCACCTCCACCGGTCCAGAGTCTTGTGCTGTCCCC 2346
|||
Db 2517 GGGCTGCGAGATGCTTCTCTATCCACCTCCACCGGTCCAGAGTCTTGTGCTGTCCCC 2576
Oy 2347 AGACCTCTGTGACACACAGCCAGATCAAGAGGACACAGGCGCAGAGATGATCTTCTTTT 2406
|||
Db 2577 AGACCTCTGTGACACACAGCCAGATCAAGAGGACACAGGCGCAGAGATGATCTTCTTTT 2636
Oy 2407 GTCTTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 2466
|||
Db 2637 GTCTTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 2696
Oy 2467 GAGAAATAAACACATTTTCAT 2488
|||
Db 2697 GAGAAATAAACACATTTTCAT 2718
|||
RESULT 8
AAH14722 standard: cDNA: 1755 BP.
ID AAH14722 standard: cDNA: 1755 BP.
AC AAH14722:
XX
DT 26-JUN-2001 (first entry)
XX
XX Human cDNA sequence SEQ ID NO:12452.
DE
KW Human; primer: detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EPI074617-A2.
XX

PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-0116126.
 XX
 PR 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 XX WPI; 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 PT
 PT
 PS Claim 8; SEQ ID 12452; 2537pp + CD ROM; English.
 PS
 XX
 CC The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC of an oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialized methods. AAH03166 to AAH13628 and
 CC AAH3633 to AAH18742 represent human cDNA sequences; AAH92446 to
 CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 XX
 SQ Sequence 1755 BP; 350 A; 577 C; 505 G; 323 T; 0 other;

Query Match 51.3%; Score 1303.6; DB 22; Length 1755;
 Best Local Similarity 94.2%; Pred. No. 8e-238;
 Matches 1430; Conservative 0; Mismatches 4; Indels 84; Gaps 4;

QY 979 ATGGGCGGGGAATTTGACAACTTTCACCTGAGCGCTTACCTACACAAATGTGGCCCTC 1038
 DB 322 ATGGCGCCGGGAATTTGACAACTTTCACCTGAGCGCTTACCTACACAAATGTGGCCCTC 381
 QY 1039 TGGATGAGAGATCGGCCAGCTGCTGCGGCACTGGAGGAGAGACCGCTTCAATTAG 1098
 DB 382 TGGATGAGAGATCGGCCAGCTGCTGCGGCACTGGAGGAGAGACCGCTTCAATTAG 441
 QY 1099 GCTGCAAGAGCACAGGGGACCCACAGTGTCTGCTCACTGCAAGATGGGCGTTCAGCGCTCA 1158
 DB 442 GCTGCAAGAGCACAGGGGACCCACAGTGTCTGCTCACTGCAAGATGGGCGTTCAGCGCTCA 501
 QY 1159 GCGGCGCAAGTGTGGGCTATGCTTATGCAAGAGTTCGATGAGCTTGGAGCAGGCGCTG 1218
 DB 502 GCGGCGCAAGTGTGGGCTATGCTTATGCAAGAGTTCGATGAGCTTGGAGCAGGCGCTG 561
 QY 1219 GCGCAGTGTGAGAGTCTCGGCGCATCGCGCCCAACCTGGCTTCTGCGCAGCTG 1278
 DB 562 GCGCAGTGTGAGAGTCTCGGCGCATCGCGCCCAACCTGGCTTCTGCGCAGCTG 621

QY 1279 CAGATCTACAGGCGCATCTGAGCGCCAGAACCTGAGGGTGTGGGAGGAGAAAGTTGT 1338
 DB 622 CAGATCTACAGGCGCATCTGAGCGCCAGAACCTGAGGGTGTGGGAGGAGAAAGTTGT 681
 QY 1339 AGGATGTGAGAGAGGCGGACGACGACCGCCGAAAGAGAGCTTGGGCGCACGCGCTATA 1398
 DB 682 AGGATGTGAGAGAGGCGGACGACGACCGCCGAAAGAGAGCTTGGGCGCAC--GCGCACCTATA 739
 QY 1399 AACCTCGAAGGGGTATGAGTTCATGATCTCTGAGACCCCTCTTGGAGCTGGAGAG 1458
 DB 740 AACCTCGAAGGGGTATGAGTTCATGATCTCTGAGACCCCTCTT--GAGCTGGAGAG 798
 QY 1459 CACCTGATGAGACATGATGATGCGAGAGTCTTCTTCCACAGAGTCTTACATGAG 1518
 DB 799 CACCTGAG--AGACCAATGATGCGAGAGTCTTCTTCCACAGAGTCTTACATGAG 857
 QY 1519 AGCCTGTGACGCGCTTCCACAGCTTGCAGAGGACCAAGGAGGCGCAGACGTGGACAGG 1578
 DB 858 AGCCTGTGACGCGCTTCCACAGCTTGCAGAGGACCAAGGAGGCGCAGACGTGGACAGG 917
 QY 1579 GGCCTCAGCCTGCGCTGAAAGTCCCGCAGTCACTGATGATACCTCCAGGCGAGTCCGTG 1638
 DB 918 GGCCTCAGCCTGCGCTGAAAGTCCCGCAGTCACTGATGATACCTCCAGGCGAGTCCGTG 977
 QY 1639 TGGCCAACCGGAGCCAGGCGCTTCCAGAGACAGAGGAGGCGCAGGCGCAGGAG 1698
 DB 978 TGGCCAACCGGAGCCAGGCGCTTCCAGAGACAGAGGAGGCGCAGGCGCAGGAG 1037
 QY 1699 AGCCTGTGATTTCTCTACAGCCAGGTTCCGGAAGTGTGTGAGACAGGCGCAGGCGCATG 1758
 DB 1038 AGCCTGTGATTTCTCTACAGCCAGGTTCCGGAAGTGTGTGAGACAGGCGCAGGCGCATG 1097
 QY 1759 ACAGTGGAGAGAGGAGGCGGAGGCTTCCAGACATGAGCCAGGCTCCCTGACATGA 1818
 DB 1098 ACAGTGGAGAGAGGAGGCGGAGGCTTCCAGACATGAGCCAGGCTCCCTGACATGA 1157
 QY 1819 AGAGATCCACAACTCTTGGAGAAACACCTCAGCTGTGTGGCGCACATCTCTCTC 1878
 DB 1158 AGAGATCCACAACTCTTGGAGAAACACCTCAGCTGTGTGGCGCACATCTCTCTC 1217
 QY 1879 AGCTCGGCGCCATACCGCTTACATGAGGCTTACCTCCAGCCCTGTCACTAGGCGCTAC 1938
 DB 1218 AGCTCGGCGCCATACCGCTTACATGAGGCTTACCTCCAGCCCTGTCACTAGGCGCTAC 1277
 QY 1939 CTCCACCCCTGTCTACTACAGCTTACCTTACAGCTTAAAGTCCAGGCGCATGTCTG 1998
 DB 1278 CTCCACCCCTGTCTACTACAGCTTACCTTACAGCTTAAAGTCCAGGCGCATGTCTG 1337
 QY 1999 CTTGTCCAGGCGCTCAGACCTTCTTAACCTGGATGTGTAGAGGAGCTGAAGTACCTTT 2058
 DB 1338 CTTGTCCAGGCGCTCAGACCTTCTTAACCTGGATGTGTAGAGGAGCTGAAGTACCTTT 1397
 QY 2059 GGGGCGAAGCAGCAGCTTATCTTCAACTTACCTTGGCAGCTTCACTGCTGTGGCAC 2118
 DB 1398 GGGGCGAAGCAGCAGCTTATCTTCAACTTACCTTGGCAGCTTCACTGCTGTGGCAC 1434
 QY 2119 GGAATGAAGAGAGCTTCCGTGCAAAAAGGATCAGGCTCCACCCCGCGCCCTTCC 2178
 DB 1435 -----CCG 1437
 QY 2179 TGCACCTCTGTCTCTCTCCAGTTCACTTCTGCAAGCAGCGCAGGCAACCACTGCG 2238
 DB 1438 TGCACCTCTGTCTCTCTCCAGTTCACTTCTGCAAGCAGCGCAGGCAACCACTGCG 1497
 QY 2239 CCCCAGAGGAGGAGATCTCTAGGCGCCAGCGCGGAGGCTGGAGAGGCTGGAGAT 2298
 DB 1498 CCCCAGAGGAGGAGATCTCTAGGCGCCAGCGCGGAGGCTGGAGAGGCTGGAGAT 1557
 QY 2299 CGCTTCCCTATCCTCAGCTCAGCGGTCCAGGCTTGTGCTGCTTCCAGAGACTTCTGTG 2358
 DB 1558 CGCTTCCCTATCCTCAGCTCAGCGGTCCAGGCTTGTGCTGCTTCCAGAGACTTCTGTG 1617
 QY 2359 ACACACAGCCAGATTCACAGGCGCACAGGCGCAGAGATAGTCTTCTTCTTCTGTG 2418

|||||
Db 1618 ACACCACGCAGATCATCAGAGGACCCAGGACAGATAGCTCTTTGCTCTTCTGSC 1677
Oy 2419 CTTCTGGCTAGTACGTTTTCATACGCTTACATATCTGCTTTGACTAGAAATAAAC 2478
Db 1678 CTTCTGGCTAGTACGTTTTCATACGCTTACATATCTGCTTTGACTAGAAATAAAC 1737
Oy 2479 ACATTTTCATATTGGTT 2496
Db 1738 ACATTTTCATATTGGTT 1755

RESULT 9
ABL40801
ID ABL40801 standard; DNA; 1755 BP.
XX
AC ABL40801;
XX
03-JUL-2002 (first entry)
XX
DE Human MAP kinase phosphatase-like enzyme DNA fragment.
XX
Mitogen activated protein; MAP; MAP kinase phosphatase-like enzyme;
KM antiasthmatic; antiidiabetic; anorectic; cytoplastic; cardiant; human;
KM antiparkinsonian; cerebroprotective; neuroprotective; nootropic; gene;
KM neuroleptic; anticonvulsant; anti-HIV; antiarrhythmic; hypotensive;
KM antiallergic; dermatological; vulnerary; gene therapy; ds.
XX
OS Homo sapiens.
XX
PN WO200220732-A2.
XX
PD 14-MAR-2002.
XX
PE 27-AUG-2001; 2001WO-EP09848.
XX
PR 07-SEP-2000; 2000US-230709P.
XX
PA (FARB) BAYER AG.
XX
PI Liou J;
XX
DR WPI; 2002-339802/37.
XX
PT New human mitogen activated protein kinase phosphatase-like enzyme
PT polypeptide, regulators of which are useful for preventing, treating
PT allergies including asthma, diabetes, obesity, cancer and
PT cardiovascular diseases -
XX
PS Disclosure; Fig 4; 134pp; English.
XX
CC The invention relates to a purified human mitogen activated protein (MAP)
CC kinase phosphatase-like enzyme polypeptide. The enzyme can be expressed
CC by standard recombinant methodology. The MAP kinase phosphatase-like
CC enzyme and encoding polynucleotides are useful for screening for
CC modulators which are used for treating a MAP kinase phosphatase-like
CC enzyme dysfunction related disease, such as asthma, a central nervous
CC system disorder, diabetes, obesity, chronic obstructive pulmonary
CC disease, cancer or a cardiovascular disease. The enzyme can be regulated
CC to treat allergies including asthma, allergic rhinitis, atopic
CC dermatitis, and anaphylaxis, central nervous system disorders such as
CC brain injuries, Parkinson's disease, dementia, multiple sclerosis,
CC stroke, Alzheimer's disease, Huntington's disease, schizophrenia, Pick's
CC disease, Creutzfeldt-Jacob dementia, progressive nuclear palsy, and human
CC immunodeficiency virus (HIV) dementia, and cardiovascular diseases
CC including myocardial infarction, ischemic diseases of the heart, atrial
CC and ventricular arrhythmia, hypertensive vascular diseases and peripheral
CC vascular diseases. The enzyme is useful in diagnostic assays for
CC detecting diseases and abnormalities or susceptibility to diseases or
CC abnormalities related to the presence of mutations in the encoding
CC nucleic acid sequences. The present sequence represents the human MAP
XX kinase phosphatase-like enzyme DNA fragment.

SQ Sequence 1755 BP; 350 A; 577 C; 505 G; 323 T; 0 other;
Query Match 51.3%; Score 1303.6; DB 24; Length 1755;
Best Local Similarity 94.2%; Pred. No. 8e-238;
Matches 1430; Conservative 0; Mismatches 4; Indels 84; Gaps 4;
Oy 979 ATGGCCCGGAGATTGACAACTTCTACCTTGAGGCTTACCTACACAAATGCGCCTC 1038
Db 322 ATGGCCCGGAGATTGACAACTTCTACCTTGAGGCTTACCTACACAAATGCGCCTC 381
Oy 1039 TGGGATGAGAGTCCGCCAGCTGCTGCCGACTGGAAGAGACGACCGCTTCAATTGAG 1098
Db 382 TGGGATGAGAGTCCGCCAGCTGCTGCCGACTGGAAGAGAGACGACCGCTTCAATTGAG 441
Oy 1099 GCTGCAAGAGCAGAGGACCCAGCTGCTGCTGCTCACTGCAATATGGGCTGACGGCTCA 1158
Db 442 GCTGCAAGAGCAGAGGACCCAGCTGCTGCTGCTCACTGCAATATGGGCTGACGGCTCA 501
Oy 1159 GCGCCACAGTGTGCGCTATGCTGCAAGCAGTACGAATGAGCGCTGAGCAGGCTCTG 1218
Db 502 GCGCCACAGTGTGCGCTATGCTGCAAGCAGTACGAATGAGCGCTGAGCAGGCTCTG 561
Oy 1219 CGCCACGTCCAGAGAGCTCCGCCATCGCCGCCAACCCCTGAGCTTCTGCGCCAGCTG 1278
Db 562 CGCCACGTCCAGAGAGCTCCGCCATCGCCGCCAACCCCTGAGCTTCTGCGCCAGCTG 621
Oy 1279 CAGATCTACCCAGGAGATCTCTGAGGCGCAGAACCTGAGGTTGGTGGGAGAGAAAGTTGT 1338
Db 622 CAGATCTACCCAGGAGATCTCTGAGGCGCAGAACCTGAGGTTGGTGGGAGAGAAAGTTGT 681
Oy 1339 AGGCATGGAAGAGAGCCAGAGCCAGCCCGAAGAAGAGCTTGGGGAGAGAGCTATA 1398
Db 682 AGGCATGGAAGAGAGCCAGAGCCAGCCCGAAGAAGAGCTTGGGGAGAGAGCTATA 739
Oy 1399 AACCTCCGAGGGGTATGAGTGCATCATGCTTCTGAGACCCCTCTTGGGAGCTGAGAG 1458
Db 740 AACCTCCGAGGGGTATGAGTGCATCATGCTTCTGAGACCCCTCTT -GAGACTGAGAG 798
Oy 1459 CAGCTCAGTAGAGCCAGTACATAGCCAGAGGCTCTCTCCAGAGCTTACATGAGAG 1518
Db 799 CAGCTCAGTAGAGCCAGTACATAGCCAGAGGCTCTCTCCAGAGCTTACATGAGAG 857
Oy 1519 AGCCTCTGAGAGCCCTTCCACAGCTTCCAGAGAGCCAGAGAGGAGGAGGAGGAG 1578
Db 858 AGCCTCTGAGAGCCCTTCCACAGCTTCCAGAGAGCCAGAGAGGAGGAGGAGGAGGAG 917
Oy 1579 GGCCTCAGCTGCTCTGAGAGTCCGCCAGTCACTGTTACCTCCAGAGGAGCTGCTGG 1638
Db 918 GGCCTCAGCTGCTCTGAGAGTCCGCCAGTCACTGTTACCTCCAGAGGAGCTGCTGG 977
Oy 1639 TGGCCACCGGAGCCAGGCTTCCAGAGAGCAGAGAGGAGGAGGAGGAGGAGGAGGAG 1698
Db 978 TGGCCACCGGAGCCAGGCTTCCAGAGAGCAGAGAGGAGGAGGAGGAGGAGGAGGAG 1037
Oy 1699 AGCCTGATTTCTCTACAGCCAGCTTCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1758
Db 1038 AGCCTGATTTCTCTACAGCCAGGCTTCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1097
Oy 1759 ACAGTGAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1818
Db 1098 ACAGTGAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1157
Oy 1819 AGAGATCCACAACTCTTGGAGAAACACCTCAGCTGCTGTTGCCGACACATTTCTCTC 1878
Db 1158 AGAGATCCACAACTCTTGGAGAAACACCTCAGCTGCTGTTGCCGACACATTTCTCTC 1217
Oy 1879 AGCTCGGCGCCCTTACCCGATCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 1938
Db 1218 AGCTCGGCGCCCTTACCCGATCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 1277
Oy 1939 CTCCACCCCTGTCTACTACAGGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 1998
Db 1278 CTCCACCCCTGTCTACTACAGGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 1337

```

QY 1999 CCTGTCCAGGGCTCAAGACTTTTCTAACTGGATGTGTAGAGGGAGTGAAGTACCTTT 2058
    |||
Db 1338 CCTTCCAGGGGCTCAAGACTTTTCTAACTGGATGTGTAGAGGGAGTGAAGTACCTTT 1337
QY 2059 GGGGGCAACAGCAGCAGCTAGTTTCTCACTCTAGCCCTGCACACTCCTGTGGCAC 2118
    |||
Db 1398 GGGGGCAACAGCAGCAGCTAGTTTCTCACTCTAG----- 1434
QY 2119 GGATGAAACAGAGCTTCCCGTCAAAAAGGATCAGCCCTCCACCCCGCCCTCCC 2178
    |||
Db 1435 -----CCC 1437
QY 2179 TGCACCTCTGTCTCTCCAGTTTCAATTCGTGAACAGCAGCCAGCAACCACTGGC 2238
    |||
Db 1438 TGCACCTCTGTCTCTCCAGTTTCAATTCGTGAACAGCAGCCAGCAACCACTGGC 1497
QY 2239 CCCCAGGAGGAGGAGATCTCTAGGCCCAAGCCGGGAGGCTGGAAGGGCTGGCAGAT 2298
    |||
Db 1498 CCCCAGGAGGAGGAGATCTCTAGGCCCAAGCCGGGAGGCTGGAAGGGCTGGCAGAT 1557
QY 2299 CGCTTCCCTCATCCACCTCCACGCGTCCAGGCTTGTGCTGTGCCAGACCTCTGTG 2358
    |||
Db 1558 CGCTTCCCTCATCCACCTCCACGCGTCCAGGCTTGTGCTGTGCCAGACCTCTGTG 1617
QY 2359 ACACGAGCCAGATCAGAGGGCAGCCAGGAGATAGTCTTTTGTCTTCTGCGC 2418
    |||
Db 1618 ACACGAGCCAGATCAGAGGGCAGCCAGGAGATAGTCTTTTGTCTTCTGCGC 1677
QY 2419 CTCGTGGTAGTACATTTTCTACCTTTACAGTATCTGCGCTTTTACTGACAATAAAC 2478
    |||
Db 1678 CTCGTGGTAGTACATTTTCTACCTTTACAGTATCTGCGCTTTTACTGACAATAAAC 1737
QY 2479 ACATTTTCATATTGTGTT 2496
    |||
Db 1738 ACATTTTCATATTGTGTT 1755

RESULT 10
ABL40805
ID ABL40805 standard; cDNA; 2322 BP.
XX
AC ABL40805;
XX
DT 03-JUL-2002 (first entry)
XX
DE Human MAP kinase phosphatase-like enzyme encoding cDNA.
XX
KW Mitogen activated protein; MAP; MAP kinase phosphatase-like enzyme;
KW antiasthmatic; antidiabetic; anorectic; cytoskeletal; cardiant; human;
KW antiparkinsonian; cerebroprotective; neuroprotective; nootropic; gene;
KW neuroleptic; anticonvulsant; anti-HIV; antiarrhythmic; hypotensive;
KW antiallergic; dermatological; vulnerary; gene therapy; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..2322
FT /*tag= a
FT /*product= "MAP kinase phosphatase-like enzyme"
FT /*note= "contains internal codon deletions"
FT /*transl_except= "(pos: 156..157, aa: Leu)"
FT /*note= "there is an apparent one codon deletion which
FT /*transl_except= "(pos: 180..181, aa: Arg)"
FT /*note= "alters the reading frame"
FT /*transl_except= "(pos: 180..181, aa: Arg)"
FT /*note= "there is an apparent one codon deletion which
FT /*transl_except= "(pos: 414..415, aa: Val)"
FT /*note= "alters the reading frame"
FT /*transl_except= "(pos: 1197..1198, aa: Tyr)"
FT /*note= "there is an apparent one codon deletion which
FT /*transl_except= "(pos: 1197..1198, aa: Tyr)"
FT /*note= "alters the reading frame"

```

```

FT /*transl_except= "(pos: 2088..2089, aa: Xaa)"
FT /*note= "there is an apparent one codon deletion which
FT /*transl_except= "(pos: 2088..2089, aa: Xaa)"
FT /*note= "alters the reading frame"
XX
PN WO200220732-A2.
XX
PD 14-MAR-2002.
XX
PF 27-AUG-2001: 2001WO-EP09848.
XX
PR 07-SEP-2000: 2000US-230709P.
XX
PA (FARB ) BAYER AG.
XX
PI Liou J:
XX
DR WPI, 2002-339802/37.
XX
DR P-PSDB; ABB07845.
XX
PT New human mitogen activated protein kinase phosphatase-like enzyme
PT polypeptide, regulators of which are useful for preventing, treating
PT allergies including asthma, diabetes, obesity, cancer and
PT cardiovascular diseases
XX
PS Claim 1; Fig 13; 134pp; English.
XX
CC The invention relates to a purified human mitogen activated protein (MAP)
CC kinase phosphatase-like enzyme polypeptide. The enzyme can be expressed
CC by standard recombinant methodology. The MAP kinase phosphatase-like
CC enzyme and encoding polynucleotides are useful for screening for
CC modulators which are used for treating a MAP kinase phosphatase-like
CC enzyme dysfunction related disease, such as asthma, a central nervous
CC system disorder, diabetes, obesity, chronic obstructive pulmonary
CC disease, cancer or a cardiovascular disease. The enzyme can be regulated
CC to treat allergies including asthma, allergic rhinitis, atopic
CC dermatitis, and anapylaxis, central nervous system disorders such as
CC brain injuries, Parkinson's disease, dementia, multiple sclerosis,
CC stroke, Alzheimer's disease, Huntington's disease, schizophrenia, Pick's
CC disease, Creutzfeldt-Jacob dementia, progressive nuclear palsy, and human
CC immunodeficiency virus (HIV) dementia, and cardiovascular diseases
CC including myocardial infarction, ischemic diseases of the heart, atrial
CC and ventricular arrhythmia, hypertensive vascular diseases and peripheral
CC vascular diseases. The enzyme is useful in diagnostic assays for
CC detecting diseases and abnormalities or susceptibility to diseases or
CC abnormalities related to the presence of mutations in the encoding
CC nucleic acid sequences. The present sequence represents the human MAP
CC kinase phosphatase-like enzyme polypeptide encoding cDNA.
XX
SQ Sequence 2322 BP; 467 A; 751 C; 718 G; 386 T; 0 other;

Query Match 50.0%; Score 1270.6; DB 24; Length 2322;
Best Local Similarity 82.0%; Pred. No. 1.5e-231;
Matches 1794; Conservative 0; Mismatches 79; Indels 316; Gaps 16;

QY 126 CGGCGCCTCCACGCCGTGGGGGCTTGAGCAGGCGGTCCAGGAAGATGCACTCCA 185
    |||
Db 147 CGCCCTCACCTGGGGCTCTCTCGCAGACGCTGTCAGGAAGATGCACTCCA 206
QY 186 GCGAAGGAGAGCTTTGCGGTGTCTC-GTGGGCTCTCTCGGAGCTGAGATGAGGGG 244
    |||
Db 207 GCGA---AGAGCTTTGCGGTGTCTC-GTGGGCTCTCTCGGAGCTGAGATGAGGGG 262
QY 245 ACAATGATGATGAGAGAGGCGAGTTCTGAGCCACAGAGAAGGCCCGAGTGAAGAG 304
    |||
Db 263 ACAATGATGATGAGAGAGGCGAGTTCTGAGCCACAGAGAAGGCCCGAGTGAAGAG 322
QY 305 AGCTTCAGGGG---ACGAGACAGACTT--CGGCAAGAGATCCACAGTCCCGAGAAGA 359
    |||
Db 323 AGCTTCAGGGGGAACCCAGACAGACTTTGCGGAGATCCAGAGATCCCGAGAGAAGA 382
QY 360 GAGAGAGAGAGGAGCAGCAGCTGACCTCATGTGACAGCTGTGAGGCCCGCAGATGACT 419
    |||
Db 383 GAGAGAGAGAGGAGCAGCAGCTGACCTCATGTGAGAGGCTGTGAGGCCCGCAGATGACT 442

```

```
QY 420 CCGCTGGCA--GCCAGCTGAGAGCACCCGGCTCCCGGCTCCGCTACCTGC----- 472
D 443 CCGACTTGGAAAGCCCAAGCTGAGAGCACCCCGGGCTCCCGGGATCCGATACCTTGCT 502
QY 473 TGTGATTTCTACACGAGAAG-GAGAGGTCTGAG-CCAGATGAGAGGCTCTCTGGG 530
D 503 TGTGATTTCTACACGAGAAGAGAGAGTGTGAGCCAGATGAGAGGCTCTCTGGG 562
QY 531 -CGTGATTTCTACAGAGAGCTCCCGAGCTGCACCTGGGCTGCTTGGCCCTCT 589
D 563 ACGTGTATTTCTTACAGAGAGCTCTCCAGCTGACCTGGGCTGCTTGGCCCTCT 622
QY 590 GGAGTGACACCCAGAGTACTTACTAGTGAAGAGGGGGCTTCAGCTGACGCTGCTGGG 649
D 623 GGAGTGACACCCAGAGTACTTACTATATGAGAGGGGGCTTCAGCTGACGCTGCTGGG 682
QY 650 AAAGCCGATCTTCAGAGCCCATCTCCATCCAGACCATGT----- 688
D 683 AAAGCCGATCTTCAGAGCCCATCTCCATCCAGACCATGTGGGCCACACTCCAGTATTGC 742
QY 689 ----- 688
D 743 ACCAAGCATGTGAGGACAGCTTAGGACAGGCTTGTACCGGGTGGCAGTGCCTACCT 802
QY 689 ----- 688
D 803 GGGCAGCGCACTACAGAGAGAGACTGAACTCCGAACAGAGCTGCCTCAATGATGAGAGG 862
QY 689 ----- 700
D 863 CTATGGCCAGCTGAGAGTCTCTGCGGCCCTCCAGCGCCAGGCTGGCGTCCAGAAC 922
QY 701 AGGACAGATGGAGAGCGCGATCCGCTGAGCTGTGAAGTGTGATGATGATGACAGC 760
D 923 AGGACAGATGGAGAGCGCGATCCGCTGAGCTGTGAAGTGTGATGATGATGACAGC 982
QY 761 TGGAGAGTGTACTTCCAAAGAGATCCCGCAGGCTCTGGAGCTGGCGCTGCCCC 820
D 983 TGGAGAGTGTGCTTCCAGAGAGATCCCGCAGGCTCTGGAGCTGGCGCTGCCCC 1042
QY 821 TCCAGCAGTACCGTGTACTTCATCGACAACAGATGCTGCTGTGGTGACAGCGGAGC 880
D 1043 TCCAGCAGTACCGTGTACTTCATCGACAACAGATGCTGCTGTGGTGACAGCGGAGC 1102
QY 881 GAGCTCCCGATCTTCCCGACCTCTACCTGAGCTGAGAGTGAAGCCAGAAACCTGG 940
D 1103 GAGCTCCCGATCTTCCCGACCTCTACCTGAGCTGAGAGTGAAGCCAGAAACCTGG 1162
QY 941 AGGAGCTGACAGAGAAAGGGTCAACCATCTTGAACATGGCCCGGAGATGACACT 1000
D 1163 AGGAGCTGACAGAGAAAGGGTCAACCATCTTGAACATGGCCCGGAGATGACACT 1219
QY 1001 TCTACCTGTAGCGCTTCACTACCAACAATGTGCGCTCTGGAGTGAAGAGTCCGAC 1060
D 1220 TCTACCTGTAGCGCTTCACTACCAACAATGTGCGCTCTGGAGTGAAGAGTCCGAC 1279
QY 1061 TGTGCGCCGATGGAAGAGAGACGACCGCTTCAATGAGGCTGCAAGAGACAGGACCC 1120
D 1280 TGTGCGCCGATGGAAGAGAGACGACCGCTTCAATGAGGCTGCAAGAGACAGGACCC 1339
QY 1121 ACGTGTGCTGCTGACAGAGTGGGCTGACGCGCTGACGCGCCAGTGTGCTGCTATG 1180
D 1340 ACGTGTGCTGCTGACAGAGTGGGCTGACGCGCTGACGCGCCAGTGTGCTGCTATG 1399
QY 1181 CCATGAACAGTACGATGACAGCTGAGAGAGGCTGCGCCAGCTGAGAGAGCTCCGCG 1240
D 1400 CCATGAACAGTACGATGACAGCTGAGAGAGGCTGCGCCAGCTGAGAGAGCTCCGCG 1459
QY 1241 CCATGCGCGCGCCCAACCTGCTGCTGCGCCAGCTGACATCTACAGGACATCCGGA 1300
D 1460 CCATGCGCGCGCCCAACCTGCTGCTGCGCCAGCTGACATCTACAGGACATCCGGA 1519
```

```
QY 1301 CG----- 1302
D 1520 CCGCAGCGCCGACAGACCATGTCTGGAGACAGAAAGTGGGTGGGCTTCGCCAGAGAGC 1579
QY 1303 -----GCCAGACTGAGGGTGTG 1322
D 1580 ACCAGCCCTGAAGTCTCTACACCATTCACACTCTTCCGCGCAGAACCTGAGGGGTGTG 1639
QY 1323 GGGAGAGAAAGTGTGAGGATGGAAGAGAGCAGGACCGCCGAAAGAGAGCTGGGG 1382
D 1640 GGGAGAGAAAGTGTGAGGATGGAAGAGAGCAGGACCGCCGAAAGAGAGCTGGGG- 1698
QY 1383 CCACGGGACGCTATTAACCTCCAGGGGCTCATGAGTTCATAGTCTTCTGGAGCCTC 1442
D 1699 -CCAGGCGCAGCTATTAACCTCCAGGGGCTCATGAGTTCATAGTCTTCTGGAGCCTC 1757
QY 1443 CTTGGAGCTGGAAGACACCTAGTAGACCACTGACATGCCAGAGAGTCTTCTCCAC 1502
D 1758 CTT-GGAGCTGGAAGACACCTAG-AGACAGTGAACATGCCAGAGGTCCTTCTCCAC 1815
QY 1503 GAGTCTTCATGAAGAGGCTGTCAGCCCTTCCACAGCTTGAAGAGCAAGGAGGCG 1562
D 1816 GAGTCTTCATGAAGAGGCTGTCAGCCCTTCCACAGCTTGAAGAGCAAGGAGGCG 1875
QY 1563 CAGCAGGTGAGACAGGGGCTCAGCCTGCTGAAGTCCCGCAGTCAAGTGTACCTTC 1622
D 1876 CAGCAGGTGAGACAGGGGCTCAGCCTGCTGAAGTCCCGCAGTCAAGTGTACCTTC 1935
QY 1623 CAGGCGAGTCCGCTGTGGGCCAACCGGACCCAGGCTTCCAGAGACAGAGAGGCGCAG 1682
D 1936 CAGGCGAGTCCGCTGTGGGCCAACCGGACCCAGGCTTCCAGAGACAGAGAGGCGCAG 1995
QY 1683 GGGCAGGGGACAGGAGAGCCCTGCAATTCCTACAGCCAGTTCGGAAGGTGGTGA 1742
D 1996 GGGCAGGGGACAGGAGAGCCCTGCAATTCCTACAGCCAGTTCGGAAGGTGGTGA 2055
QY 1743 CAGGCGAGGCTCATGACAGTGAAGAGAGGCGGAGGCTGAGCCTCAGATGCCAC 1802
D 2056 CAGGCGAGGCTCATGACAGTGAAGAGAGGCGGAGGCTGAGCCTCAGATGCCAC 2113
QY 1803 GCTCCCTGACACTGAAGAGATCCAACTCTTGGAGAAACACCTTCAGCTGTGTC 1862
D 2114 GCTCCCTGACACTGAAGAGATCCAACTCTTGGAGAAACACCTTCAGCTGTGTC 2173
QY 1863 GGCACAACTCTCTACAGTCCGCCCCCATACCGGTCACAGCTCAGCTCCACCT 1922
D 2174 GGCACAACTCTCTACAGTCCGCCCCCATACCGGTCACAGCTCAGCTCCACCT 2233
QY 1923 GTCACAGGCGCTCACTCCACCGCCCTGACACTACAGGCTCAGCTCAGCTTAAGT 1982
D 2234 GTCACAGGCGCTCACTCCACCGCCCTGACACTACAGGCTCAGCTCAGCTTAAGT 2293
QY 1983 CCCAGGCCCATGTGCTGCTGTCCAAAGGCG 2011
D 2294 CCCAGGCCCATGTGCTGCTGTCCAAAGGCG 2322

RESULT 11
ABN59832
ID ABN59832 standard; cDNA: 2061 BP.
XX
XX ABN59832;
XX
XX 28-JUN-2002 (first entry)
XX
XX Novel human coding sequence SEQ ID NO: 243.
XX
XX Human; antianaemic; vulnereary; antiinflammatory; immunomodulator;
XX antineoplastic; cerebroprotective; cytostatic; rheumatic; gene therapy;
XX neuroprotective; antiparkinsonian; protein therapy; EST;
XX expressed sequence tag; gene; ss.
XX
XX Homo sapiens.
```

XX WO200222660-A2.
PN 21-MAR-2002.
XX 10-SEP-2001; 2001WO-US26015.
XX 11-SEP-2000; 2000US-0659671.
XX (HYSEQ-) HYSEQ INC.
XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AD, Yang Y, Wehrman T, Drmanac RT;
XX MPI: 2002-292408/33.
DR P-PSDB: ABB97419.
XX
XX An isolated polynucleotide for treating diseases associated with its
PT encoded polypeptide such as cancer and multiple sclerosis.
XX
XX Claim 1; SEQ ID NO 243; 509bp; English.
PS
CC The present invention provides the protein and coding sequences of 444
CC novel human proteins. These were isolated from expressed sequences tags
CC (ESTs). They can be used to stimulate cell growth, to regulate
CC hematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
CC e.g. in burn treatment, to regulate the immune system e.g. to treat
CC multiple sclerosis, to regulate actin or thymolysis e.g. to treat
CC infertility, to regulate haemostasis or thrombolysis e.g. to treat
CC stroke and cancer, to screen for drugs, to treat inflammatory conditions
CC e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.
CC Parkinson's disease. The present sequence is a coding sequence of the
CC invention.
XX
SQ Sequence 2061 BP; 415 A; 672 C; 605 G; 369 T; 0 other:

Query Match 47.9%; Score 1216.4; DB 24; Length 2061;
Best Local Similarity 74.6%; Pred. No. 2.8e-221;
Matches 1875; Conservative 0; Mismatches 126; Indels 513; Gaps 7;

QY 1 CCGTGTCTCCGGGCTCCAGGACTGTCCCGGGGTTGAGGGAAGGGCCGCTCCCGGTGC 60
DB 61 CCGTGTCTCCGGGCTCCAGGACTGT-CCGCGGGGTTGAGGGAAGGGCCGCTCCCGGTGC 119
QY 61 CAGCCAGGTGCTCGCGGCTGCTCATGCGCTGTGTCACAGTGAAGCCGTTGCGCCCG 120
DB 120 CAGCCAGGTGCTCGCGGCTGCTCATGCGCTGTGTCACAGTGAAGCCGTTGCGCCCG 179
QY 121 GGCAGCGGCGCTCCAGCGCCGCTGGGGCCCTGGACACAGGCGGTCCAGGAAGAGTGA 180
DB 180 GGCAGCGGCGCTCCAGCGCCGCTGGGGCCCTGGACACAGGCGGTCCAGGAAGAGTGA 239
QY 181 CTCGACGCAAGGACAGAGCTTTCGCTGCTCGTGGGGCTGCTCTGGAGTGAAGTGA 240
DB 240 CTCGACGCAAGGACAGAGCTTTCGCTGCTCGTGGGGCTGCTCTGGAGTGAAGTGA 299
QY 241 GGGGCAATGATGATGACAGAGGCGCAAGTCTGAGCCAAAGAGAGGCGCCGATGAG 300
DB 300 GGGGCAATGATGATGACAGAGGCGCAAGTCTGAGCCAAAGAGAGGCGCCGATGAG 359
QY 301 GAGAGCTTCACGGGGACACAGACTTGGGCAAGAGTCCAGAGTCCCAAGAAAGCAG 360
DB 360 GAGAGCTTCACGGGGACACAGACTTGGGCAAGAGTCCAGAGTCCCAAGAAAGCAG 419
QY 361 GAGGAGCAGAGGACACCTGACCTCATGATGATGAGTCTGAGGCGGAGATGACATC 420
DB 420 GAGGAGCAGAGGACACCTGACCTCATGATGATGAGTCTGAGGCGGAGATGACATC 479
QY 421 CCGCTGGAGACCCAGCTGAGAGCACCAGGCGCTCCCGGCTCCGCTACCTGCTGATGTT 480
DB 480 CCGCTGGAGACCCAGCTGAGAGCACCAGGCGCTCCCGGCTCCGCTACCTGCTGATGTT 539
QY 481 TCTACACGAAAGAGAGGTCTGAGCCAGATGAGACGGTCTCTGCGGCTGATTTTC 540

DB 540 TCTACACGAAAGAGAGGTCTGAGCCAGATGAGAGGCTCTCTGGGGTGGATTTTC 599
QY 541 CCTGACAGCAGCTCCCGCAGCTGACACCTGGGCGCTGCTTGGCCCTGAGATGACACC 600
DB 600 CCTGACAGCAGCTCCCGCAGCTGACACCTGGGCGCTGCTTGGCCCTGAGATGACACC 659
QY 601 CAGGTGTACTAGATGAGAGCGGGGCTTCAGCTGACGTCTGTTGGCCAAAGCCGATC 660
DB 660 CAGGTGTACTAGATGAGAGCGGGGCTTCAGCTGACGTCTGTTGGCCAAAGCCGATC 719
QY 661 TTCAAGCCCATCTCCATCCAGACCATGTGCTCCAGAAACAGAGCAGATGAGAGCGG 720
DB 720 TTCAAGCCCATCTCCATCCAGACCATGTGCTCCAGACCATCTCCAGG----- 762
QY 721 ATCCGTCTGAGCTGTGAAAGTGTGATGTCAGTGCATCTGAGAGTGTACTTCCAAA 780
DB 763 ----- 762
QY 781 GAGATCCCGCAGGCTCTGAGCTGCGCTGGGGCTCCCTCCAGACGATACCGTATTC 840
DB 783 -----TATTGC 768
QY 841 ATGCAACAACAGATGCTGCTGCTGTCGACACAGCGGACCGAGCCCTCCCGCATCTCC 900
DB 769 ACCAAGCATGTGAGGACCTCTAGGACGCGCTTGTACCGGG----- 811
QY 901 CACCTCTACCTGGCTCAGAGTGAAGCAGCAACCTGAGAGCTGACAGAGAACAG 960
DB 812 ----- 811
QY 961 GTACCCACATCTTAAACATGAGCCCGGAGATTGACAACTTACCTGAGGCTTACAC 1020
DB 812 -----TGCGAGTGCCTCACTCAGCTGGGCGAGCAGCATCAGAGAGACTGAA 857
QY 1021 TACCAATGTCGCGCTCGGATGAGAGTGGGCCAGCTCTCCGACATGAGAAAGAG 1080
DB 858 TCCGAAACAGAGTGCCTCAATGA----- 880
QY 1081 ACGCAGCGCTTCAATTGAGGCTCAAGAGCAGAGGCGCAGCAGCTGTGCTCACTGCAAG 1140
DB 881 ----- 880
QY 1141 ATGGGCTGAGCGCTCAGCGGCGCAGAGTGTGCTGCTATGCAATGAAAGCATACGATGC 1200
DB 881 ----- 880
QY 1201 AGCTGAGACAGCCCTGCGCCAGCGAGAGCTCCGCGCATGCGCCCGCAACCT 1260
DB 881 -----GTGACGCGCTATGCGCCAGCTGAGTCTCTGGGCTCCAGCGCGCAGCTGGC 935
QY 1261 GCGTTCCTCGCCAGCTGAGATCTACAGGGCATCTGACGGCCAGAACTGAGGGTGG 1320
DB 936 GCGTTCCTCAGAACAGAGAGAGATGAGAGCGAGGATC-----GTGC 976
QY 1321 TGGGAGAGAGAGGTTGTAGGCTGAGAAAGACCAAGGAGCGCCCAAGAAAGAGCTGG 1380
DB 977 TGAGCTGTGAGAAAGT----- 991
QY 1381 GGCACAGGGGCAAGTAAACCTCCGAGGGGTATGAGTCCATGATCTTCTGAGAGCC 1440
DB 992 ----- 991
QY 1441 TCCCTGGAGCTGAGAGCACCCTAGTAGACAGTACATGACAGAGTCTTCTCTCC 1500
DB 992 -----GTGAGAGCTGAGAGAGACCTCAG-AGACAGTATACATGCAAGAGTCTTCTCTCC 1047
QY 1501 ACGAGTCTTACATGAGAGCTCTGACGCTTCCACAGCTTGCAGAGACCAAGGAG 1560
DB 1048 ACGAGTCTTACATGAGAGCTCTGACGCTTCCACAGCTTGCAGAGACCAAGGAG 1107
QY 1561 GCCAGAGTGTGACAGAGGGGCTCAGGCTGAGAGTCCGCGCAGTACAGGTTACCC 1620

Db 1108 GCCACGAGTGTGACAGAGGGGCGCTCAGCTGCCCTGAAAGTCCCGCCAGTCACTGATTACCC 1167
 Qy 1621 TCACAGGAGTATCCCTGTGTGGCCCAACCGAGCCCTTCACAGAGCAGAGAGAGGCG 1680
 Db 1168 TCACAGGAGTATCCCTGTGTGGCCCAACCGAGCCCTTCACAGAGCAGAGAGAGGCG 1227
 Qy 1681 AGGGGAGGGGAGGAGAGAGCCCTGATTTCTCTACGCCAGTTCGGAAGTGTGTGA 1740
 Db 1228 AGGGGAGGGGAGGAGAGAGCCCTGATTTCTCTACGCCAGTTCGGAAGTGTGTGA 1287
 Qy 1741 GACAGAGCCAGGCTGATATCACTATGAGAGAGGCGAGGCGCTGAGCCCTCACACTGCC 1800
 Db 1288 GACAGAGCCAGGCTGATATCACTATGAGAGAGGCGAGGCGCTGAGCCCTCACACTGCC 1347
 Qy 1801 AGGCTCCCTGTGACATGAAAGAGATCCAACTCCTTGGAGAAACACCTCAGCTGTGT 1860
 Db 1348 AGGCTCCCTGTGACATGAAAGAGATCCAACTCCTTGGAGAAACACCTCAGCTGTGT 1407
 Qy 1861 GCCGACACATTTCTCTGAGTCCGCGCCCATATCCGCTACTACAGCCCTCAGCTCCACCC 1920
 Db 1408 GCCGACACATTTCTCTGAGTCCGCGCCCATATCCGCTACTACAGCCCTCAGCTCCACCC 1467
 Qy 1921 CTGTACTACAGGCTCAGCTCCACCCCTGTCTACTACAGCCCTCAGCTCCCTTAA 1980
 Db 1468 CTGTACTACAGGCTCAGCTCCACCCCTGTCTACTACAGCCCTCAGCTCCCTTAA 1527
 Qy 1981 GTCCAGGCGCCATGTCTGCTGCTCCAGAGGCTCAAGACTTTTAACTGGAGTGTGTGA 2040
 Db 1528 GTCCAGGCGCCATGTCTGCTGCTCCAGAGGCTCAAGACTTTTAACTGGAGTGTGTGA 1587
 Qy 2041 GGGACTGAAGTACTCTTTGGGGGCAACAGCACCCTAGTTTCTTCAACTTACCCCTG 2100
 Db 1588 GGGACTGAAGTACTCTTTGGGGGCAACAGCACCCTAGTTTCTTCAACTTACCCCTG 1647
 Qy 2101 CACACTCAGCTGTGGCAGCGAATGAAAACAGACTTCCCGTGCAAAAAGGCTCAGCCCTC 2160
 Db 1648 CACACTCAGCTGTGGCAGCGAATGAAAACAGACTTCCCGTGCAAAAAGGCTCAGCCCTC 1707
 Qy 2161 CCACCCCGCCCGCCCTCTGACACCTCTGCTCTCCAGTTCATCTCGAAGCAGGCA 2220
 Db 1708 CCACCCCGCCCGCCCTCTGACACCTCTGCTCTCCAGTTCATCTCGAAGCAGGCA 1767
 Qy 2221 GGGCAGGAGACCGAGTGGCCCGCAAAAGCAGGAGGATCTCAGGCCCGCAGCCGGGAG 2280
 Db 1768 GGGCAGGAGACCGAGTGGCCCGCAAAAGCAGGAGGATCTCAGGCCCGCAGCCGGGAG 1827
 Qy 2281 CTGGAAGGGCTGGCAGATCGCTTCCCTCATCTCCACCGGCTCAGGCTCTTTCTGCT 2340
 Db 1828 CTGGAAGGGCTGGCAGATCGCTTCCCTCATCTCCACCGGCTCAGGCTCTTTCTGCT 1887
 Qy 2341 GTCCCGACAGCTCTGTGACACAGCCAGATCACAGGGGACACAGGCCAGAGATAGTCTT 2400
 Db 1888 GTCCCGACAGCTCTGTGACACAGCCAGATCACAGGGGACACAGGCCAGAGATAGTCTT 1947
 Qy 2401 CTTTTCCTCTTTCTGGGCTCTGAGCTAGTCAATTTTTCATAGCCCTTACAGTATGAGCTT 2460
 Db 1948 CTTTTCCTCTTTCTGGGCTCTGAGCTAGTCAATTTTTCATAGCCCTTACAGTATGAGCTT 2007
 Qy 2461 TGTACTGAGAAATAAACAATTTTCATATTTGTTAAAAA 2514
 Db 2008 TGTACTGAGAAATAAACAATTTTCATATTTGTTAAAAA 2061
 RESULT 12
 AAF63578 standard; cDNA: 1026 BP.
 AAF63578;
 AC AAF63578;
 XX 11-MAY-2001 (first entry)
 DF Human phosphatase NP_060746_h coding sequence.
 DE
 XX

KW Phosphatase; gene therapy; enzyme; cancer; pathophysiological hypoxia;
 KW cardiac dysfunction; vascular disorder; myopathy; ectodermal dysplasia;
 KW congenital muscle disorder; Papillon-Leleuve syndrome; Cowden disease;
 KW Moebius syndrome; Bjornstad syndrome; Bannayan Zonana syndrome;
 KW schizophrenia; hamartoma; ss.
 OS Homo sapiens.
 PN MO200112819-A2.
 PD 22-FEB-2001.
 PF 11-AUG-2000; 2000WO-US22158.
 PR 13-AUG-1999; 99US-0149005.
 PA (SUGEN-) SUGEN INC.
 PI Plowman GD, Martinez R, Whyte D, Hill R, Flanagan P, Lioubin M;
 DR WPI: 2001-211226/21.
 DR P-PSDB: AAB73226.
 XX
 PT New protein phosphatase polypeptide for diagnosing and treating
 PT phosphatase related disorders such as cancer, schizophrenia, cardiac
 PT dysfunction and/or vascular disorders -
 XX
 PS Claim 1: Fig 4; 138pp; English.
 CC
 CC The present invention relates to phosphatase proteins and coding
 CC sequences. The present sequence is one such phosphatase coding sequence.
 CC Phosphatases are enzymes that catalyse the dephosphorylation of proteins
 CC modified by phosphorylation of serine, threonine or tyrosine residues.
 CC The phosphatases are useful for treating a variety of diseases: for
 CC example cancer e.g. breast, urogenital, prostate, head, neck, lung
 CC cancers, synovial sarcomas, renal cell carcinoma, non-small cell lung
 CC cancer, hepatocellular carcinoma, pancreatic endocrine tumours, stomach
 CC cancer, glioblastoma, colorectal cancer and thyroid cancer,
 CC pathophysiological hypoxia, cardiac dysfunction and/or vascular
 CC disorders, myopathies, congenital muscle disorders, Papillon-Leleuve
 CC syndrome, Cowden disease, ectodermal dysplasia, Moebius syndrome,
 CC Bjornstad syndrome, Bannayan Zonana syndrome, schizophrenia and
 CC hamartomas.
 SQ Sequence 1026 BP: 217 A; 324 C; 326 G; 159 T; 0 other:
 Query Match 30.7%; Score 779.6; DB 22; Length 1026;
 Best Local Similarity 89.7%; Pred. No. 1.5e-138;
 Matches 924; Conservative 0; Mismatches 4; Indels 102; Gaps 4;
 Qy 853 ATGCTGCTGTGTGGCAGCGGAGCGGAGCCGATCTTCCGACCTCTACCTG 912
 Db 1 ATGCTGCTGTGTGGCAGCGGAGCGGAGCCGATCTTCCGACCTCTACCTG 60
 Qy 913 GGCTCAGAGTGGAGCAGCAAACTGGAGAGCTGACAGAGAACAGGCTACCCACATC 972
 Db 61 GGCTCAGAGTGGAGCAGCAAACTGGAGAGCTGACAGAGAACAGGCTACCCACATC 120
 Qy 973 TTGAACATGGCCCGGAGATTGACAACTTTCTACCGTGAAGCTTACCTACCAATGTG 1032
 Db 121 TTGAACATGGCCCGGAGATTGACAACTTTCTACCGTGAAGCTTACCTACCAATGTG 180
 Qy 1033 GCGCTGTGGAGTGAAGAGTGGCGGAGCTGCTGGCGACTGGAAGAGAGCGACGGCTTC 1092
 Db 181 GCGCTGTGGAGTGAAGAGTGGCGGAGCTGCTGGCGACTGGAAGAGAGCGACGGCTTC 240
 Qy 1093 ATTGAGGCTGCAAGAGCAGAGGACCCACAGTGTGTGTCCATGCAAGATGGCGCTCAGC 1152
 Db 241 ATTGAGGCTGCAAGAGCAGAGGACCCACAGTGTGTGTCCATGCAAGATGGCGCTCAGC 300
 Qy 1153 CGCTCAGCGGCGACAGTGTGCTATGCAATGAAGCAGTACGAATGACGCTGGAGCAG 1212
 Db 301 CGCTCAGCGGCGACAGTGTGCTATGCAATGAAGCAGTACGAATGACGCTGGAGCAG 360

```
QY 1213 GCCCTGCGCAGCTGAGAGCTCCGGCCCATCGCCGCCCAACCTGGCTTCCTGCGC 1272
    |||||||
    361 GCCCTGGCCACGTCGAGAGCTCGGGCCCATCGCCGCCCAACCTGGCTTCCTGCGC 420
QY 1273 CAGCTGCAGATCTACCGAGGCTCTGACG----- 1302
    |||||||
    421 CAGCTGCAGATCTACCGAGGCTCTGACGCGCCAGCCGCCAGACCATGCTGGGAGCAG 480
QY 1303 ----- 1302
    |||||||
    481 AAGTGGGTGGGCTCTCCAGAGAGGACACCACCCCTGAAGTCTTACACCATTTCCCA 540
QY 1303 -----GCCAGAACCTGAGAGGTGTGGGAGAGAGAGGTTTGGCATGAGAGAGAGC 1354
    |||||||
    541 CCTCTCCGCCAGAACCTGAGAGGTGTGGGAGAGAGAGGTTTGGCATGAGAGAGAGC 600
QY 1355 CAGCAGACCCCGAAAGAGAGCTGGGGCCACGGGGCAGCTATTAACCTCCGAGGGTCA 1414
    |||||||
    601 CAGCAGACCCCGAAAGAGAGCTGGGGCCAC--GGCCACGTAATAACCTCCGAGGGTCA 658
QY 1415 TGAGTGCATGATCTTGAGGCTCTGAGGCTCTTGGAGGCTGGAGAGACCTGACATGAGACAG 1474
    |||||||
    659 TGAGTGCATGATCTTGAGGCTCTTCTCCACAGAGTCTTCAATGAAAGCTCTGACGCCCTT 716
QY 1475 TGACATGCGAGAGGTCTTCTCTTCCACAGAGTCTTCAATGAAAGCTCTGACGCCCTT 1534
    |||||||
    717 TGACATGCGAGAGGTCTTCTCTTCCACAGAGTCTTCAATGAAAGCTCTGACGCCCTT 776
QY 1535 CCCACAGCTTTCAGAGACCAAGGAGGCGCAGAGTGGACAGAGGGGCTTACGCTGCCCT 1594
    |||||||
    777 CCCACAGCTTTCAGAGACCAAGGAGGCGCAGAGTGGACAGAGGGGCTTACGCTGCCCT 836
QY 1595 GAATCCCGCCAGTCAGTGTACCTTCACAGGAGTGGCTGGTGGCCAAACCGGAGCCA 1654
    |||||||
    837 GAATCCCGCCAGTCAGTGTACCTTCACAGGAGTGGCTGGTGGCCAAACCGGAGCCA 896
QY 1655 GGCCTTCAGAGAGCAGAGAGCAGGGGCGAGGGGCGAGGAGAGAGCCCTGATTTCTC 1714
    |||||||
    897 GGCCTTCAGAGAGCAGAGAGCAGGGGCGAGGGGCGAGGAGAGAGCCCTGATTTCTC 956
QY 1715 TACGCCCAAGTTCGGGAAGTGTGAGACAGGCGCAGCTGATACAGTGGAGAGAGGG 1774
    |||||||
    957 TACGCCCAAGTTCGGGAAGTGTGAGACAGGCGCAGCTGATACAGTGGAGAGAGGG 1016
QY 1775 CGAGGCTCGA 1784
    |||||||
    1017 CGAGGCTCGA 1026
Db
QY 1775 CGAGGCTCGA 1784
    |||||||
    1017 CGAGGCTCGA 1026
Db
RESULT 13
ABL40803
ID ABL40803 standard; DNA: 599 BP.
XX
AC ABL40803:
XX
DT 03-JUL-2002 (first entry)
XX
DE Human MAP kinase phosphatase-like enzyme DNA fragment.
XX
KW Mitogen activated protein; MAP; MAP kinase phosphatase-like enzyme;
KW antihistaminic; antidiabetic; anorectic; cytosolic; cardiac; human;
KW antiparkinsonian; cerebroprotective; neuroprotective; nootropic; gene;
KW neuroleptic; anticonvulsant; anti-HIV; antiarthritic; hypotensive;
KW antiallergic; dermatological; vulnery; gene therapy; ds.
XX
OS Homo sapiens.
XX
PN WO200220732-A2.
XX
PD 14-MAR-2002.
XX
PF 27-AUG-2001; 2001WO-EP09848.
```

```
XX 07-SEP-2000; 2000US-230709P.
PR (FARB ) BAYER AG.
XX
PA
XX
PI Lhou J;
XX
DR WPI; 2002-339802/37.
XX
PT New human mitogen activated protein kinase phosphatase-like enzyme
PT polypeptide, regulators of which are useful for preventing, treating
PT allergies including asthma, diabetes, obesity, cancer and
PT cardiovascular diseases
XX
XX Disclosure; Fig 8; 134pp; English.
XX
XX CC The invention relates to a purified human mitogen activated protein (MAP)
CC kinase phosphatase-like enzyme polypeptide. The enzyme can be expressed
CC by standard recombinant methodology. The MAP kinase phosphatase-like
CC enzyme and encoding polynucleotides are useful for screening for
CC modulators which are used for treating a MAP kinase phosphatase-like
CC enzyme dysfunction related disease, such as asthma, a central nervous
CC system disorder, diabetes, obesity, chronic obstructive pulmonary
CC disease, cancer or a cardiovascular disease. The enzyme can be regulated
CC to treat allergies including asthma, allergic rhinitis, atopic
CC dermatitis, and anaphylaxis, central nervous system disorders such as
CC brain injuries, Parkinson's disease, dementia, multiple sclerosis,
CC stroke, Alzheimer's disease, Huntington's disease, schizophrenia, Pick's
CC disease, Creutzfeldt-Jacob dementia, progressive nuclear palsy, and human
CC immunodeficiency virus (HIV) dementia, and cardiovascular diseases
CC including myocardial infarction, ischemic diseases of the heart, atrial
CC and ventricular arrhythmia, hypertensive vascular diseases and peripheral
CC vascular diseases. The enzyme is useful in diagnostic assays for
CC detecting diseases and abnormalities or susceptibility to diseases or
CC abnormalities related to the presence of mutations in the encoding
CC nucleic acid sequences. The present sequence represents the human MAP
CC kinase phosphatase-like enzyme DNA fragment.
XX
SQ Sequence 599 BP; 135 A; 185 C; 183 G; 96 T; 0 other;
XX
Query Match 22.0%; Score 558.6; DB 24; Length 599;
Best Local Similarity 98.7%; Pred. No. 1,1e-96;
Matches 595; Conservative 0; Mismatches 4; Indels 4; Gaps 3;
QY 922 TGGAGCGACCAACCTGAGAGAGTGCAGAGAGAGAGGTCACCCACATCTTGAACATG 981
    |||||||
    1 TGGAGCGACCAACCTGAGAGAGTGCAGAGAGAGAGGTCACCCACATCTTGAACATG 60
Db
QY 982 GCCCGGAGATTGACAACTTCTACCTGAGCGCTTACCTACCAATGTCGCGCTTGG 1041
    |||||||
    61 GCCCGGAGATTGACAACTTCTACCTGAGCGCTTACCTACCAATGTCGCGCTTGG 120
QY 1042 GATGAGAGTGGGCCAGAGTGTGCGGCACTGGAAGAGAGCGAGCGTCAATGAGGCT 1101
    |||||||
    121 GATGAGAGTGGGCCAGAGTGTGCGGCACTGGAAGAGAGCGAGCGTCAATGAGGCT 180
Db
QY 1102 GCAAGAGACAGGGGACCCACAGTGTGTCACACTGCAAGTGGGCGCTCAGCGCTAGCG 1161
    |||||||
    181 GCAAGAGACAGGGGACCCACAGTGTGTCACACTGCAAGTGGGCGCTCAGCGCTAGCG 240
QY 1162 GCCACAGTGTGCGCTATGCGATGAAGACAGTACGAATGACAGCTTGAAGAGGCGCTTGGCC 1221
    |||||||
    241 GCCACAGTGTGCGCTATGCGATGAAGACAGTACGAATGACAGCTTGAAGAGGCGCTTGGCC 300
Db
QY 1222 CAGGTGAGAGTTCGGGCCATGCGCCGCCAGCCAGCCCTGCTCTGCGCAGCTGCGAG 1281
    |||||||
    301 CAGGTGAGAGTTCGGGCCATGCGCCGCCAGCCAGCCCTGCTCTGCGCAGCTGCGAG 360
QY 1282 ATCTACAGGCGATCTCTGAGGCGCAGAACTGAGGTGTGGGAGAGAGAGGTTTGAAG 1341
    |||||||
    361 ATCTACAGGCGATCTCTGAGGCGCAGAACTGAGGTGTGGGAGAGAGAGGTTTGAAG 420
QY 1342 CATGAAGAGAGCGCAGGACGCCCGGAAAGAGAGCTGGGGGCCACGGGGCAGCTATAAC 1401
```

|||||
DB 421 CATGAGAGAGACCCAGCAGAGCCCCGAAAGAGAGCTGGG--CCACGGCCAGCATTAAC 478
OY 1402 CTCGAGGGGTCATAGAGTCCATCATGCTCTTCGGAGCCCTCTTGGAGCTGGAGAGC 1461
DB 479 CTCGAGGGGTCATAGAGTCCATCATGCTCTTCGGAGCCCTCTT--GGAGCTGGAGAGC 537
OY 1462 CTCAGTAACACGAGTACATGACAGAGTCTTCTCTCCACAGATCTTCAATGAAGAC 1521
DB 538 CTCAG-AGACGATGACATGCCAGAGTCTTCTCTCCACAGATCTTCAATGAAGAC 596
OY 1522 CTC 1524
DB 597 CTC 599
RESULT 14
ABL40802
ID ABL40802 standard: DNA; 409 BP.
AC ABL40802;
XX
DT 03-JUL-2002 (first entry)
DE Human MAP kinase phosphatase-like enzyme DNA fragment.
XX
KW Mitogen activated protein; MAP; MAP kinase phosphatase-like enzyme;
KW antiasthmatic; antidiabetic; anorectic; cyrostatic; cardiant; human;
KW antiparkinsonian; cerebroprotective; neuroprotective; nootropic; gene;
KW neuroleptic; anticonvulsant; anti-HIV; antiarrhythmic; hypotensive;
KW antifallergic; dermatological; vulnary; gene therapy; ds.
XX
OS Homo sapiens.
XX
PN WO200220732-A2.
PD 14-MAR-2002.
PF 27-AUG-2001; 2001WO-EP09848.
PR 07-SEP-2000; 2000US-230709P.
XX
PA (FARB) BAYER AG.
XX
PI Liou J;
XX
DR WPI: 2002-339802/37.
XX
PT New human mitogen activated protein kinase phosphatase-like enzyme
PT polypeptide, regulators of which are useful for preventing, treating
PT allergies including asthma, diabetes, obesity, cancer and
PT cardiovascular diseases
XX
PS Disclosure: Fig 7; 134pp; English.
XX
CC The invention relates to a purified human mitogen activated protein (MAP)
CC kinase phosphatase-like enzyme polypeptide. The enzyme can be expressed
CC by standard recombinant methodology. The MAP kinase phosphatase-like
CC enzyme and encoding polynucleotides are useful for screening for
CC modulators which are used for treating a MAP kinase phosphatase-like
CC enzyme dysfunction related disease, such as asthma, a central nervous
CC system disorder, diabetes, obesity, chronic obstructive pulmonary
CC disease, cancer or a cardiovascular disease. The enzyme can be regulated
CC to treat allergies including asthma, allergic rhinitis, atopic
CC dermatitis, and anaphylaxis, central nervous system disorders such as
CC brain injuries, Parkinson's disease, dementia, multiple sclerosis,
CC stroke, Alzheimer's disease, Huntington's disease, schizophrenia, Pick's
CC disease, Creutzfeldt-Jacob dementia, progressive nuclear palsy, and human
CC immunodeficiency virus (HIV) dementia, and cardiovascular diseases
CC including myocardial infarction, ischaemic diseases of the heart, atrial
CC and ventricular arrhythmia, hypertensive vascular diseases and peripheral
CC vascular diseases. The enzyme is useful in diagnostic assays for
CC detecting diseases and abnormalities or susceptibility to diseases or

CC abnormalities related to the presence of mutations in the encoding
CC nucleic acid sequences. The present sequence represents the human MAP
CC kinase phosphatase-like enzyme DNA fragment.
XX
SQ Sequence 409 BP; 87 A; 127 C; 121 G; 73 T; 1 other:
Query Match 15.6%; Score 397; DB 24; Length 409;
Best Local Similarity 99.5%; Pred. No. 4,2e-66;
Matches 408; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
OY 717 GCGATCCGCTGCTGAGCTGTGGAAGTGTGATGTCACTGAGAGTGCATCTT 776
DB 1 GCGATCCGCTGCT-AGCTGTGGAAGTGTGATGTCACTGAGAGTGCATCTT 59
OY 777 CAAAGATCCGCGCAGGCTGTGAGCTGTGCGGCTGCCCTCCACAGTACCGTGA 836
DB 60 CAAGAGATCCGCGCAGGCTGTGAGCTGTGCGGCTGCCCTCCACAGTACCGTGA 119
OY 837 CTTTCATCGACAACAGATGCTCTGCTGTGGCACAGCGGACCGCTCCCGCATCTT 896
DB 120 CTTTCATCGACAACAGATGCTCTGCTGTGGCACAGCGGACCGCTCCCGCATCTT 179
OY 897 CCCCCACCTCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 956
DB 180 CCCCCACCTCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 239
OY 957 CAGGGTCACCCACATCTTGAACATGGCCCGGAGATGTGCAACTTCTACCGAGGCTT 1016
DB 240 CAGGGTCACCCACATCTTGAACATGGCCCGGAGATGTGCAACTTCTACCGAGGCTT 299
OY 1017 CACCTACCAACATGTGCGCTCTGGATGAGAGTGTGCGGCCAGCTGCTGCCACTGGA 1076
DB 300 CACCTACCAACATGTGCGCTCTGGATGAGAGTGTGCGGCCAGCTGCTGCCACTGGA 359
OY 1077 GGAGAGCGACCCGCTTCAATGAGCGCTGCAAGAGCACAGGCGCACCGCTGC 1126
DB 360 GGAGAGCGACCCGCTTCAATGAGCGCTGCAAGAGCACAGGCGCACCGCTGC 409
RESULT 15
AAH98183/C
ID AAH98183 standard: cDNA; 717 BP.
AC AAH98183;
XX
PI 12-OCT-2001 (first entry)
XX
DE Human EST-derived coding sequence SEQ ID NO: 40.
XX
KW Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
KW diagnostics; forensic test; gene mapping; genetic disorder;
KW biodiversity; gene therapy; nutrition; ss.
XX
OS Homo sapiens.
XX
PN WO200154477-A2.
PD 02-AUG-2001.
PF 25-JAN-2001; 2001WO-US02687.
PR 25-JAN-2000; 2000US-0491404.
PR 17-JUL-2000; 2000US-0617746.
PR 03-AUG-2000; 2000US-0631451.
PR 15-SEP-2000; 2000US-0663870.
XX
PA (HXSE-) HXSEQ INC.
XX
PI Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
PI Cao Y, Dmanac RA, Zhang J, Werhman T;
XX
DR WPI: 2001-476164/51.

Db 629 GTCTACTGCAAGAACGGCCGACGCCGCTCGGCCGCCGCTGCAACCGCGTACCTCATGGG 688

QY 1189 CAGTACGAATGACGCTGGAGCAGGCGCCGACCTGCGCAGAGAGCTCCGCGCCATCGCC 1248
DB 689 CACCGCGGCGCTGAGCGGGAAGGCTTCAGATGGTGAAGAGCCCTCGCCCGGAGCA 748
QY 1249 CCGCCCAACCCCTGCTTCTGCGCCGAGCTGCGATCTACCGGCACTCTGACGGCCACA 1308
DB 749 GAACCCAGACCCGCGCTTCTGCTGCTCAGCTCCAGAACTAGAGAGGCGCTCCAGGCGCCAG 808
QY 1309 ACCTGAGGCTGGTGGGAGAGAGAGTTGATGAGCATGGAAGAGCCAGGAGCC 1362
DB 809 TCCTGCTGCGAGGAGAGAGCCCGCCTTAGGCTTGGGCGCTGAGGCTTGAAGC 862

RESULT 2
US-09-704-139-3
; Sequence 3, Application US/09704139
; Patent No. 6420153
; GENERAL INFORMATION:
; APPLICANT: Kapeller, Rosana
; TITLE OF INVENTION: 18232, A NOVEL DUAL SPECIFICITY PHOSPHATASE AND USES THEREFOR
; FILE REFERENCE: 10448-018001
; CURRENT APPLICATION NUMBER: US/09/704,139
; PRIORITY FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: US 60/185,772
; FILING DATE: 2000-02-29
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-704-139-3

Query Match 4.1%; Score 103.4; DB 4; Length 531;
Best Local Similarity 52.7%; Pred. No. 5.3e-14;
Matches 224; Conservative 0; Mismatches 201; Indels 0; Gaps 0;

QY 889 CGCATTTTCCCGACCTCTACTGAGGCTGAGTGAAGCAGCAAACTGGAGAGCTG 948
DB 61 CGCGTGGCGCTCACTCTCTCTGCGAGCGCGGAGCGCGGCGGAGGAGCAGCTG 120
QY 949 CAGAGAACAGGCTCAACCATCTTGAAATGAGGCGCGGAGATTTACAATCTTACCT 1008
DB 121 GCGCGCGGAGAGTCAAGCTGTGCTTCCCGCAGAGAGCCGCGCGCGCGCG 180
QY 1009 GAGCGCTTCACTTACCAATGAGCTGTGGATGAGAGTGGCGCCAGCTGCGG 1068
DB 181 CCGCGCTGCGAGAGTGGCGCGCTGTGCTGAGAGAGCCGCTGAGAGCTGCTGCG 240
QY 1069 CACTGAGAGAGCAGCAGCTTCAATTGAGGCTGCAAGAGCAGAGGCAAGCCAGTGTG 1128
DB 241 CACTGAGAGCAGCAGCTGCGCGCGCAGAGAGCGCGGCGCGCGCGCTGCTA 300
QY 1129 GTCCACTGGAAGTGGCGCTGAGCGCTGAGCGGCGCAGCTGTGCTATGCAATGAAG 1188
DB 301 GTTCTACTGGAAGAGCGCGCAGCGCTGCGCGCTGTGCAAGCGCTGATCTGAGCG 360
QY 1189 CAGTACGAATGACGCTGGAGCAGGCGCTGCGCAGCTGAGAGAGCTCCGCGCCATCGCC 1248
DB 361 CACCGGCGCTGAGCTGCGGAGAGGCTTCCAGATGGTGAAGAGCGCTCGCCGGTAGCA 420
QY 1249 CCGCCCAACCCCTGCTTCTGCGCGCAGCTGCAAGTGTACAGAGGCAATCTGAGCGCCACA 1308
DB 421 GAACCCAGACCCGCGCTTCTGCTCAGCTCCAGAACTAGAGAGGCGCTCCAGGCGCCAG 480
QY 1309 ACCTG 1313
DB 481 TCCTG 485

RESULT 3
US-08-530-290-11
; Sequence 11, Application US/08530290

; Patent No. 5958721
; GENERAL INFORMATION:
; APPLICANT: Marshall, Christopher John
; APPLICANT: Ashworth, Alan
; TITLE OF INVENTION: Hughes, David Anthony
; TITLE OF INVENTION: Methods for Screening of Substances for
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/530,290
; FILING DATE: 14-DEC-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/GB94/00694
; FILING DATE: 31-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9402573.1
; FILING DATE: 10-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9307250.2
; FILING DATE: 07-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 084611-000000US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1238 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-530-290-11

Query Match 3.5%; Score 88; DB 2; Length 1238;
Best Local Similarity 50.5%; Pred. No. 2e-10;
Matches 214; Conservative 0; Mismatches 210; Indels 0; Gaps 0;

QY 892 ATCTTCCCGACCTTACTGAGGCTCAGAGTGAAGCAGCAAACTGGAGAGCTGAG 951
DB 648 ATCTTCCCGACCTTACTGAGGCTCAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 707
QY 952 AGCAAGAGGCTACCCACATCTTGAACATGAGCGCGGAGATTGACAACTTACCTGAG 1011
DB 708 GCGCTGGGATACAGGCTCTGTGTAATGTCTCTGAGACTGCCCAACCATCTTGAAGA 767
QY 1012 CGCTTACCTTACCAATGAGTGGCGCTTGGGATGAAGAGTGGCCAGCTGCTGCGCAC 1071
DB 768 CACTATCATGATACAGTATCCCAATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 827
QY 1072 TGGAGAGAGAGCAGCCCTTCAATTGAGGCTGCAAGAGCAGCAGGCAAGCCAGCTGCTG 1131
DB 828 TTCAATGAGAGCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 887
QY 1132 CACTGCAAGATGGGCTCAGCGGCTCAGGCGCAGAGTGGGCTTATGCAATGAGAGG 1191
DB 888 CACTGCAAGGCGGATCTCGCGGTGCGCAGCAGCATCTGCTGCTTACCTGATGATGAG 947

NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/045,973
FILING DATE: Filed Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
ATTORNEY/AGENT INFORMATION:
FILING DATE:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0491 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1729 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRSTN016
CLONE: 3041794
US-09-045-973-6

Query Match 2.4%; Score 61.6; DB 4; Length 1729;
Best Local Similarity 46.8%; Pred. No. 0.00016;
Matches 193; Conservative 0; Mismatches 219; Indels 0; Gaps 0;
QY 887 CCGGCACTTCCCGCCCTTACCTGGGCTCAGAGTGGAAAGCAGCAACCTGAGGAGC 946
DB 574 CTCAAATCACCTCTCTCTATCTCTGGGAGAGGCAAGTGGCTTCAATCGCACCTCC 633
QY 947 TGGAGAGGAACAGGTCACCATCTGTAACATGGCCGGAGATTGACACTTCTACC 1006
DB 634 TCCAGGCTCTGGCATCAGCTGATTTATGCTTACATTGAGATCCCTAATTTCACT 693
QY 1007 CTGAGCGCTTACCTACACATATGCGCCTCTGGAGTGAAGAGTGGCCAGCTGCTGC 1066
DB 694 GGGCCCAATTGAGTAAAGTGGCTGCTGGCTGACATGGCCGCAATGGCCCATGGAG 753
QY 1067 CGCAGCTGAAGAGACCCACCGCTTCAATTAGGCTGCAAGACAGCAGGCAACCACTGC 1126
DB 754 TGTACTTGGACACCGCTGCTGACAGATCCAGATGTGTGACAGAGACAGGAGCCACT 813
QY 1127 TGGTCCACTGCAGATGGGGGTGACGGCTCAGAGGGCCACAGTGGCTGTATGGCATGA 1186
DB 814 TGGGCACTGTGCTGACGGGGGTGAGCCGCTCAGCAGCAGCTGTGTATGCTTACCTGTA 873
QY 1187 AGCAGTACGATGAGCCTGAGAGAGGCTGCGCCAGCTGAGAGAGCTCCGGCCCATCG 1246
DB 874 AATTCCAAAGCTGTGCTGCTGAGAGCGCTACAACTGGGTAAGAAAGCCCGGAGCTGTCA 933
QY 1247 CCGGCCCCAACCTGCTCTCTCGCCAGCTGAGATCTACAGGAGCATCT 1298
DB 934 TCAAGCCCAAGTAGGCTCTGAGAGCACTGATAGACTACAGAGCCGCACT 985

RESULT 7

US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PTZgpt-F1s
US-08-232-463-14

Query Match 2.4%; Score 60; DB 1; Length 7218;
Best Local Similarity 2.4%; Pred. No. 0.00068;
Matches 9; Conservative 229; Mismatches 144; Indels 0; Gaps 0;
QY 1835 CTGTGAGAAACCCCTCAGCTGTGCGCAGACATTCCTCTCAGTCCGCCCATACC 1894
DB 1059 CTTCGATATTT 1118
QY 1895 CGTCACTACAGCCTCAGCTCCACCCCTGTCACTAGCGCTCAGCTCCACCCCTGTAC 1954
DB 1119 YY 1178
QY 1955 TACAGCTCAGCTCCAGCTTAAGTCCAGGCCAGGCAAGTGTGCTGTCCAAAGGCTCA 2014
DB 1179 YY 1238
QY 2015 AGACTTCTACAGGAGTGTGAGAGGAGTGAAGTACCTTTGGGGGCAACAGCACCC 2074
DB 1239 YY 1298
QY 2075 TACTTATCTCACTCAGCCTGACACTCAGCTGTGCGAGGAAAGAAAGAGAGC 2134
DB 1299 YY 1358
QY 2135 TTCCGCTGCAAAAAGGCTCAGCGCTCCACCCCGCCCTCTGACACTCTCTCTCTC 2194

LENGTH: 991 base pairs

```

;
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; TISSUE TYPE: SOYBEAN
; IMMEDIATE SOURCE:
; CLONE: SSM.PK0067.G5
;
US-09-296-715-25

```

```

Query Match      2.3%; Score 57.4; DB 4; Length 991;
Best Local Similarity 76.9%; Pred. No. 0.0011;
Matches 70; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

```

```

QY 2450 GTATCTGGCTTGTACTGAGAAATAAACACATTTTCATATTTGGTTAAAAA 2509
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 884 GTATGCTATTTTAACTTAACTAAAAAGGTCAGTTTAAAAA 943
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 2510 AAAAAAAAAAAAAAAAAAAAAAAAAA 2540
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 944 AAAAAAAAAAAAAAAAAAAAAAAAAA 974
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

Search completed: April 11, 2003, 23:54:49
Job time : 183.246 secs

```

QY	1	CTGTGTCCTGCGGGGTCCAGAGACTGTCCCGGGGGTTCAGGAGAAAGGGCCGTGGCCGGTGC	60
Db	1	CTGTGTCCTGCGGGGTCCAGAGACTGTCCCGGGGGTTCAGGAGAAAGGGCCGTGGCCGGTGC	60
QY	61	CAGCCAGAGTCTCGCGGGCTGGGCTCCATGTGCCCTCTGTGCACAGTAGAGCCGTTTGCCTCCG	120
Db	61	CAGCCAGAGTCTCGCGGGCTGGGCTCCATGTGCCCTCTGTGCACAGTAGAGCCGTTTGCCTCCG	120
QY	121	GGCAGCGGCGCTCCACGCCCGGTGGGGCCCTTGGGACACAGGGGGGTTCAGCGCAAGAGATCGA	180
Db	121	GGCAGCGGCGCTCCACGCCCGGTGGGGCCCTTGGGACACAGGGGGGTTCAGCGCAAGAGATCGA	180
QY	181	CTCCAGCGCAAGGAGAGCTTTCGCGGTGTCCGTGGGAGCTGTCTCCGTGGAGCTCCACAGATGGA	240
Db	181	CTCCAGCGCAAGGAGAGCTTTCGCGGTGTCCGTGGGAGCTGTCTCCGTGGAGCTCCACAGATGGA	240
QY	241	GGGACCAATGATGATGCAGCAGAGAGCCAGTTCTAGCCACAGAGAAAGGCCCGAGATGAG	300

Db 241 GGGGACATGATGATGACGACGAGGCCAGTTCTGAGCCAAAGAGAGGCCCGGAGTGA 300
QY 301 GAGAGCTCCACGGGGACGACAGACTTGGGGAAGATCCAGAGTCCCGAAGACG 360
Db 301 GAGAGCTCCACGGGGACGACAGACTTGGGGAAGATCCCGAAGACG 360
QY 361 GAGAGCAGAGGACGACGCTGACACCTCATGATGATGAGCTGCTGAGGCCGAGATGACATC 420
Db 361 GAGAGCAGAGGACGACGCTGACACCTCATGATGATGAGCTGCTGAGGCCGAGATGACATC 420
QY 421 CGCTGGCAGCCGACGCTGAGGAGGACCCCGGCTCCCGGCTCCGCTACCTGCTGATGTT 480
Db 421 CGCTGGCAGCCGACGCTGAGGAGGACCCCGGCTCCCGGCTCCGCTACCTGCTGATGTT 480
QY 481 TCTACAGAGAGAGAGAGTCTGAGCCAGATGAGAGGCTCTCTCGGCGCGATGTT 540
Db 481 TCTACAGAGAGAGAGAGTCTGAGCCAGATGAGAGGCTCTCTCGGCGCGATGTT 540
QY 541 CTTGACAGAGCTCCCGACGCTGACCCCTGGCTGCTTGGCTTGGCTGAGTGAAC 600
Db 541 CTTGACAGAGCTCCCGACGCTGACCCCTGGCTGCTTGGCTTGGCTGAGTGAAC 600
QY 601 CAGGTACTTAAATGAGAGCGGGGCTTGAAGTGAAGCTGCTGAGGAGAAACCGGATC 660
Db 601 CAGGTACTTAAATGAGAGCGGGGCTTGAAGTGAAGCTGCTGAGGAGAAACCGGATC 660
QY 661 TTCAAGCCATCTTCATCCAGACCATGCTGCTCAGAAAGAGAGAGAGATGAGAGAGCG 720
Db 661 TTCAAGCCATCTTCATCCAGACCATGCTGCTCAGAAAGAGAGAGAGATGAGAGAGCG 720
QY 721 ATCCGCTGAGCTGCTGAGAAAGTGTGATGTCAGTGAACCTGAGAGTGTACTTCCAA 780
Db 721 ATCCGCTGAGCTGCTGAGAAAGTGTGATGTCAGTGAACCTGAGAGTGTACTTCCAA 780
QY 781 GAGATCCGAGGCTGAGAGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 840
Db 781 GAGATCCGAGGCTGAGAGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 840
QY 841 ATGACAAACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Db 841 ATGACAAACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
QY 901 CACTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Db 901 CACTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
QY 961 GTCAACCAATCTTGAACATGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 1020
Db 961 GTCAACCAATCTTGAACATGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 1020
QY 1021 TACCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
Db 1021 TACCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
QY 1081 ACGCAGCGCTTCAATGAGCTGCAAGAGCAGAGGCAACCCAGCTGCTGCTGCTGCTG 1140
Db 1081 ACGCAGCGCTTCAATGAGCTGCAAGAGCAGAGGCAACCCAGCTGCTGCTGCTGCTG 1140
QY 1141 ATGGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 1200
Db 1141 ATGGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 1200
QY 1201 AGCTGAGAGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 1260
Db 1201 AGCTGAGAGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 1260
QY 1261 GCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
Db 1261 GCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
QY 1321 TGGGAGAGAGAGGCTTGTAGCATGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
Db 1321 TGGGAGAGAGAGGCTTGTAGCATGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380

QY 1381 GGGCAGGGGACAGTAAACCTCCGAGGGGCTCATGAGGTCCATGCTTCTGAGGCC 1440
Db 1381 GGGCAGGGGACAGTAAACCTCCGAGGGGCTCATGAGGTCCATGCTTCTGAGGCC 1440
QY 1441 TCCCTTGGAGCTGAGAGACCTCAGTAGACAGTAGACATGCGAAGGTCTTCTTCCC 1500
Db 1441 TCCCTTGGAGCTGAGAGACCTCAGTAGACAGTAGACATGCGAAGGTCTTCTTCCC 1500
QY 1501 ACGAGCTTCAATGAGAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGG 1560
Db 1501 ACGAGCTTCAATGAGAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGG 1560
QY 1561 GCGAGAGGTGAGACAGGGGGCTCAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 1620
Db 1561 GCGAGAGGTGAGACAGGGGGCTCAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 1620
QY 1621 TCCAGGAGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680
Db 1621 TCCAGGAGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680
QY 1681 AAGGAGAGAGAGAGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740
Db 1681 AAGGAGAGAGAGAGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740
QY 1741 GACAGGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800
Db 1741 GACAGGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800
QY 1801 ACGCTTCCCTGACACTGAAGAGATCCACACTCTCTTGAAGAACCTCAGCTGCTT 1860
Db 1801 ACGCTTCCCTGACACTGAAGAGATCCACACTCTCTTGAAGAACCTCAGCTGCTT 1860
QY 1861 GCGGACACATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
Db 1861 GCGGACACATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
QY 1921 CTGTACTAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1980
Db 1921 CTGTACTAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1980
QY 1981 GTCCAGGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040
Db 1981 GTCCAGGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040
QY 2041 GGGAGTGAAGTACCTTGGGGGCAACAGACCTAGTTCATTCATTCATTCATTCATTC 2100
Db 2041 GGGAGTGAAGTACCTTGGGGGCAACAGACCTAGTTCATTCATTCATTCATTCATTC 2100
QY 2101 CACACTCACCTGTGCGACGGAATGAAGACAGAGCTTCCGTCAGAAAGAGGTCAACGCT 2160
Db 2101 CACACTCACCTGTGCGACGGAATGAAGACAGAGCTTCCGTCAGAAAGAGGTCAACGCT 2160
QY 2161 CCAACCCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2220
Db 2161 CCAACCCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2220
QY 2221 GGGCAGAGCAACAGTGGGCGCCCAAGAGAGAGAGTCCAGGCGCCGAGGCGGAGG 2280
Db 2221 GGGCAGAGCAACAGTGGGCGCCCAAGAGAGAGAGTCCAGGCGCCGAGGCGGAGG 2280
QY 2281 CTGGAAGGGCTGGAGATGCTTCCCTCATTCACCTCCACCGGTCAGAGTCTTCTGCT 2340
Db 2281 CTGGAAGGGCTGGAGATGCTTCCCTCATTCACCTCCACCGGTCAGAGTCTTCTGCT 2340
QY 2341 GTCCCAAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400
Db 2341 GTCCCAAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400
QY 2401 CTCTTGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460
Db 2401 CTCTTGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460

OY	2461	TGCTCAGAAATAAAAACATTTTCATATTGTGTAATAAAAAAAAAA	2520
Db	2461	TGTCTCAGAAATAAAACCATTTTTTCAATTGGTTAAAAA	2520
OY	2521	AAAAAAAAAAAAAAAAA	2540
b	2521	AAAAAAAAAAAAAAAAA	2540

RESULT 2

```

US-09-761-640-1
: Sequence 1, Application US/09761640
: Patent No US20020137042A1
:
: GENERAL INFORMATION:
:
: APPLICANT: WEI, Ming-Hui et al
:
: TITLE OF INVENTION: ISOLATED HUMAN PHOSPHATASE PROTEINS,
: TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHATASE PROTEINS,
: TITLE OF INVENTION: AND USES THEREOF
:
: FILE REFERENCE: CL000964-CIP
:
: CURRENT APPLICATION NUMBER: US/09/761,640
:
: CURRENT FILING DATE: 2001-01-18
:
: NUMBER OF SEQ ID NOS: 10
:
: SOFTWARE: FastSeq for Windows Version 4.0
:
: SEQ ID NO 1
:
: LENGTH: 2704
:
: TYPE: DNA
:
: ORGANISM: Human
:
: US-09-761-640-1

```

Query Match	88.4%;	Score 2246.4;	DB 10;	Length 2704;
Best Local Similarity	92.4%;	Pred. No. 0;		
Matches 2497;	Conservative 0;	Mismatches 11;	Indels 194;	Gaps 5;

QY	1	CTGTGATCTCTGCGGATCCAGAGC	181	CTCCCGCGGGGGG11GAGAGAGAGGGG	323	GTGCGGAGC	60
Db	8	CTGTGCTCTCGGGGATCCAGAGACTGT	-CCGCGGGGGTTGAGGGAGAGGGGGCGTGGCCGGGTGC	66			
QY	61	CAGCCCAAGTGTCTGCGGGGCTTGAGCTCCATGAGCCCTTGATGACAGTTCAGCCGTTCCGGCCCCG	120				
Db	67	CAGGCCAAGTGTCTGCGGGCTTGAGCTCCATGAGCCCTGTGTACAGTGAAGCGTTCCGCCCCCG	126				
QY	121	GGCAGCGGGGCTCTCACGCCCTGTGGGGCCCTGGGACCAAGCGGCTGCACAGGAGAGTTCGA	180				
Db	127	GGCAGCGGGGCTCTCACGCCCTGTGGGGCCCTGGGACCAAGCGGCTGCACAGGAGAGTTCGA	186				
QY	181	CTTCAGGAGGAGGAGAGACTTGTGGGGGTCTCCGTTGGGGGGCTCTCTGGAGCTGACAGATGGA	240				
Db	187	CTTCAGGAGGAGGAGAGACTTGTGGGGGTCTCCGTTGGGGGGCTCTCTGGAGCTGACAGATGGA	246				
QY	241	GGGGACAATGATGATGCAGACAGAGAGGCCAGTTCTGAGCCCAAGAGAAAGGCCCGCGAGTGA	300				
Db	247	GGGGACAATGATGATGCAGACAGAGAGGCCAGTTCTGAGCCCAAGAGAAAGGCCCGCGAGTGA	306				
QY	301	GAGAGACTCCACGGGGGACCCAGACAGACTTTCGGGGCAAGGATCCCAAGATGCCCAAGACAG	360				
Db	307	GAGAGCTTCACGGGGGACCCAGACAGACTTTCGGGGCAAGGATCCCAAGATGCCCAAGACAG	366				
QY	361	GAGAGCAGAGAGGAGACCTGCACACTCATGTGTACAAGCTCTAGGCGCCAGAGATGACATC	420				
Db	367	GAGAGCAGAGAGGAGACCTGCACACTCATGTGTACAAGCTCTAGGCGCCAGAGATGACATC	426				
QY	421	CGCCTGGCAGCCAGCTGGAGGACACCCGGGCTCCCGGGCTCCGCTACCTGCTGGTAGTT	480				
Db	427	CGCCTGGCAGCCAGCTGGAGGACACCCGGGCTCCCGGGCTCCGCTACCTGCTGGTAGTT	486				
QY	481	TCTACAGAGAAAGAGAAAGTCTGAGCCAGAGATGAGAGGTCCTCTGGGGGTGGATTTC	540				
Db	487	TCTACAGAGAAAGAGAAAGTCTGAGCCAGAGATGAGAGCGTCTCTTGGGGGTGGATTTC	546				
QY	541	CTGTACAGACGCTCCCGCAGCTGCACCTGTGGGCTGTGTTCCCTCCCTCTGGAGTGAACCC	600				
Db	547	CTGTACAGACGCTCCCGCAGCTGCACCTGTGGGCTGTGTTCCCTCCCTCTGGAGTGAACCC	606				

Db 727 GAGGAGCTTAGGAGGCGCCTTGTACCGGGTGGCAGTGCCTCCTCAGCTGGCCAGCCAG 786
 Oy 689 ----- 688
 Db 787 TACGAGAGAGACTAAGTCCGACAGAGCTGCTCAATGATGAGCGCTTATGGCCAG 846
 Oy 689 -----GGTCTCAGAAACAGGAGGAGATG 711
 Db 847 CTGAGTCTCTGGGCGCTCCAGCCGCGAGCCTGGCGGCTCTCAGAAACAGGAGGAGATG 906
 Oy 712 GAGCAGGCGATCCGTGCTGAGCTGTGGAAGTGTGATGTGATGATGACCTGGAGAGTGTG 771
 Db 907 GAGCAGGCGATCCGTGCTGAGCTGTGGAAGTGTGATGTGATGATGACCTGGAGAGTGTG 966
 Oy 772 ACTTCCAAAGAGATCCGCGAGGCTCTGAGGCTGGCGCTGGGGCTCCCTCCAGAGATGAG 831
 Db 967 ACTTCCAAAGAGATCCGCGAGGCTCTGAGGCTGGCGCTGGGGCTCCCTCCAGAGATGAG 1026
 Oy 832 CGTGACTTCATGAGACAGACAGATGCTGCTGTGTGGCAGACGGGAGCCGAGCTCCGCG 891
 Db 1027 CGTGACTTCATGAGACAGACAGATGCTGCTGTGTGGCAGACGGGAGCCGAGCTCCGCG 1086
 Oy 892 ATCTTCCCGCCACCTTCTACCTGGGGCTGAGAGTGAAGCGACAGCAACCTGGAGAGCTGAG 951
 Db 1087 ATCTTCCCGCCACCTTCTACCTGGGGCTGAGAGTGAAGCGACAGCAACCTGGAGAGCTGAG 1146
 Oy 952 AGGAACAGGCTACCCACATCTTGAACATGGCCCGGAGATGAGCAACTTCAACCTGAG 1011
 Db 1147 AGGAACAGGCTACCCACATCTTGAACATGGCCCGGAGATGAGCAACTTCAACCTGAG 1206
 Oy 1012 CGCTTACCTACACAGATGCTGCGCTGTGGATGAGAGTGGCCAGCTGCTGGCGCAC 1071
 Db 1207 CGCTTACCTACACAGATGCTGCGCTGTGGATGAGAGTGGCCAGCTGCTGGCGCAC 1266
 Oy 1072 TGGAAAGGAGACCCACCGCTTCAATTGAGCTGCAAGAGCAGAGGCAACCCAGCTGCTG 1131
 Db 1267 TGGAAAGGAGACCCACCGCTTCAATTGAGCTGCAAGAGCAGAGGCAACCCAGCTGCTG 1326
 Oy 1132 CAGTCAAGATGGGGGTGACGGCGCTGAGGGCGCACAGTCTGGCTATGCTCAATGAAAGCG 1191
 Db 1327 CAGTCAAGATGGGGGTGACGGCGCTGAGGGCGCACAGTCTGGCTATGCTCAATGAAAGCG 1386
 Oy 1192 TACGAATGAGAGCTGAGAGAGGCGCTGCGCCAGCTGACAGAGAGCTCCGGCCATTCGCGCC 1251
 Db 1387 TACGAATGAGAGCTGAGAGAGGCGCTGCGCCAGCTGACAGAGAGCTCCGGCCATTCGCGCC 1446
 Oy 1252 CCCAACCTCGGCTCTCTCGCGCAGCTGACAGATCTACAGAGGATCTCTGAGCGCCAGAAC 1311
 Db 1447 CCCAACCTCGGCTCTCTCGCGCAGCTGACAGATCTACAGAGGATCTCTGAGCGCCAGAAC 1506
 Oy 1312 TGAAGGTGGTGGGAGAGAGTGTGAGGATGGAAGAGAGCCAGGCGCCGCAAGA 1371
 Db 1507 TGAAGGTGGTGGGAGAGAGTGTGAGGATGGAAGAGAGCCAGGCGCCGCAAGA 1566
 Oy 1372 AGAGCTGGGGCAGAGGCGCAGTATAAAGCTCCGAGGGGTCATGAGTCCATCACTT 1431
 Db 1567 AGAGCTGGG--CAAGGCGCAGTATAAAGCTCCGAGGGGTCATGAGTCCATCACTT 1624
 Oy 1432 CTGAGAGCTCTCTGGAGCTGGAAGAGCAGCTAGTAGACAGTGAATGCCAGAGGCTCT 1491
 Db 1625 CTGAGAGCTCTCT--GAGAGCTGAGAGCAGCTCAG--AGACAGTGAATGCCAGAGGCTCT 1682
 Oy 1492 TCTCTTCCAGAGTCTCAGATGAAGAGCCTTGAGAGCTCTCCACAGCTTGAAGA 1551
 Db 1683 TCTCTTCCAGAGTCTCAGATGAAGAGCCTTGAGAGCTCTCCACAGCTTGAAGA 1742
 Oy 1552 CCAGAGGAGCCAGAGTGAAGAGGGGGCTCAGCTGCTGGAAGTCCCGCAGATCAG 1611
 Db 1743 CCAGAGGAGCCAGAGTGAAGAGGGGGCTCAGCTGCTGGAAGTCCCGCAGATCAG 1802
 Oy 1612 TGGTTAACCTCCAGGAGCAGTGGCTGTGTGGCCAGCCGAGCCAGGCTTCCAGAGAGCAG 1671
 Db 1803 TGGTTAACCTCCAGGAGCAGTGGCTGTGTGGCCAGCCGAGCCAGGCTTCCAGAGAGCAG 1862

Oy 1672 AGCAGGGGAGGGGAGGGGAGGAGAGCCCTGCAATTTCTCTAGAGCCAGATTCCGA 1731
 Db 1863 AGCAGGGGAGGGGAGGGGAGGAGAGCCCTGCAATTTCTCTAGAGCCAGATTCCGA 1922
 Oy 1732 AGTGGGTGAGAGAGGCGCAGCTGATGACAGTGAAGAGAGGGGAGGCGCTGAGCCCTCA 1791
 Db 1923 AGTGGGTGAGAGAGGCGCAGCTGATGACAGTGAAGAGAGGGGAGGCGCTGAGCCCTCA 1982
 Oy 1792 CACATGCCAGGCTCCCTGACACTGGAAGAGATCCACACTTCCTTGGGAAACACCTTC 1851
 Db 1983 CACATGCCAGGCTCCCTGACACTGGAAGAGATCCACACTTCCTTGGGAAACACCTTC 2042
 Oy 1852 AGCTGTGGGCGCAGACATTCCTCTCAGCTCGGCGCCATACCCGTCATGAGCTTCAC 1911
 Db 2043 AGCTGTGGGCGCAGACATTCCTCTCAGCTCGGCGCCATACCCGTCATGAGCTTCAC 2102
 Oy 1912 CTCCACACCCCTGTCACTACAGGCGCTCAGCTCCACCCCTGTCACTACAGCTCAGCTCA 1971
 Db 2103 CTCCACACCCCTGTCACTACAGGCGCTCAGCTCCACCCCTGTCACTACAGCTCAGCTCA 2162
 Oy 1972 CAGCTTAAAGTCCAGAGCCATGTCTGCTGTCCAAAGGCTCAAGACTTCTAACTGGGA 2031
 Db 2163 CAGCTTAAAGTCCAGAGCCATGTCTGCTGTCCAAAGGCTCAAGACTTCTAACTGGGA 2222
 Oy 2032 TGTGTAGAGGAGTGAAGATACCTTTGGGGGCAACAGCAGCTAGTTTCAATTCAGACT 2091
 Db 2223 TGTGTAGAGGAGTGAAGATACCTTTGGGGGCAACAGCAGCTAGTTTCAATTCAGACT 2282
 Oy 2092 CTAGCCCTGACACACTCACCTGTGGGACGGAATGAAGAACAGACTTCCCGTGAAGAGGG 2151
 Db 2283 CTAGCCCTGACACACTCACCTGTGGGACGGAATGAAGAACAGACTTCCCGTGAAGAGGG 2342
 Oy 2152 TTAGGCTTCCACCCCGGCGCTGCTGACCTGCTGTCTCTGCTGCTGCTGCTGCTGCTG 2211
 Db 2343 TTAGGCTTCCACCCCGGCGCTGCTGACCTGCTGTCTCTGCTGCTGCTGCTGCTGCTG 2402
 Oy 2212 AACGAGCCAGGCGAGGCAACAGTGGAGCCGCAAGGAGGAGAGTCCAGAGCCAGC 2271
 Db 2403 AACGAGCCAGGCGAGGCAACAGTGGAGCCGCAAGGAGGAGAGTCCAGAGCCAGC 2462
 Oy 2272 CGGAGGAGGCTGGAAGGCGTGGCAGATCGCTTCCATCAGCTCCACCGGTCAGGTC 2331
 Db 2463 CGGAGGAGGCTGGAAGGCGTGGCAGATCGCTTCCATCAGCTCCACCGGTCAGGTC 2522
 Oy 2332 TTTGCTGCTGCTCCAGAGACTCTGTGAGACAGCGCAGATCAAGAGGAGCAGGCGCAGA 2391
 Db 2523 TTTGCTGCTGCTCCAGAGACTCTGTGAGACAGCGCAGATCAAGAGGAGCAGGCGCAGA 2582
 Oy 2392 GATAGCTCTCTTTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2451
 Db 2583 GATAGCTCTCTTTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2642
 Oy 2452 ATCTGGCTTTGCTAGTGAAGATTAACACATTTTCATATTTGTTAAAAA 2511
 Db 2643 ATCTGGCTTTGCTAGTGAAGATTAACACATTTTCATATTTGTTAAAAA 2702
 Oy 2512 AA 2513
 Db 2703 AA 2704

RESULT 4
 US-09-761-640-2
 ; Sequence 2, Application US/09761640
 ; Patent No. US20020137042A1
 ; GENERAL INFORMATION:
 ; APPLICANT: WEI, Ming-Hui et al
 ; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHATASE PROTEINS,
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHATASE PROTEINS,
 ; FILE REFERENCE: CLO00964-CIP
 ; CURRENT APPLICATION NUMBER: US/09/761,640

```
; CURRENT FILING DATE: 2001-01-18
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 2852
; TYPE: DNA
; ORGANISM: Human
US-09-761-640-2

Query Match      84.5%  Score 2145.4  DB 10; Length 2852;
Best Local Similarity 89.2%  Pred. No. 0;
Matches 2493; Conservative 0; Mismatches 11; Indels 291; Gaps 5;

QY      33  GGTGAGGAGAGGGGCGCTGCCCCGCTGCCAGCCAGGTGCTCCGGCTTGCTCATATGCG 92
Db      2  GGTGAGGAGAGGGGCGCTGCCCCGCTGCCAGCCAGGTGCTCCGGCTTGCTCATATGCG 61

QY      93  CCTGTGTCACAGTAGACCGCTTCCGCCGGGACGGCGCCTCCAGCGCCGTGGGGCCCTG 152
Db      62  CCTGTGTCACAGTAGACCGCTTCCGCCGGGACGGCGCCTCCAGCGCCGTGGGGCCCTG 121

QY      153  GGAACAGAGGGGCTGACAGAGAGTAGTCCAGGAGGACAGAGGACAGGCTTGGGGTCTCCG 212
Db      122  GGAACAGAGGGGCTGACAGAGAGTAGTCCAGGAGGACAGAGGACAGGCTTGGGGTCTCCG 181

QY      213  TGGGGCTGTCTTGGGACTGCGAGGATGGAGGGGCAATATGATGACAGAGGCCAGTTTC 272
Db      182  TGGGGCTGTCTTGGGACTGCGAGGATGGAGGGGCAATATGATGACAGAGGCCAGTTTC 241

QY      273  TGAGCCACACAGAGAGAGGCCCCGAGTAGAGAGAGCTCCAGGGGACACAGAGACTTCGG 332
Db      242  TGAGCCACACAGAGAGAGGCCCCGAGTAGAGAGAGCTCCAGGGGACACAGAGACTTCGG 301

QY      333  GCAGAGATCCAGAGTCCCGAGAGAGAGAGAGAGAGGACAGGACAGCTGACCTCATGTGT 392
Db      302  GCAGAGATCCAGAGTCCCGAGAGAGAGAGAGAGAGGACAGGACAGCTGACCTCATGTGT 361

QY      393  ACAGCTGTGAGAGCCGACAGGATGACATCCGCTGGGACAGCCAGCTGGAGGACCCGGCC 452
Db      362  ACAGCTGTGAGAGCCGACAGGATGACATCCGCTGGGACAGCCAGCTGGAGGACCCGGCC 421

QY      453  TCCCGGGCTCCGCTACCTGCTGTAGTTTCTACACAGAGAGAGAGAGGTGTGAGCCAGGA 512
Db      422  TCCCGGGCTCCGCTACCTGCTGTAGTTTCTACACAGAGAGAGAGGTGTGAGCCAGGA 481

QY      513  TGAGAGCGTCTCTCTGGGGGTGATTTCCCTGACAGAGAGTCCCGCAGCTGCACCTGGG 572
Db      482  TGAGAGCGTCTCTCTGGGGGTGATTTCCCTGACAGAGAGTCCCGCAGCTGCACCTGGG 541

QY      573  CCGTGTCTTGGCCCTCTGAGTAGACACCCAGGTGTACTTATAGATGAGAGCGGGGCTTCAG 632
Db      542  CCGTGTCTTGGCCCTCTGAGTAGACACCCAGGTGTACTTATAGATGAGAGCGGGGCTTCAG 601

QY      633  CGTAGAGCTGTGGTGGGCAAGCGGATCTTCAAGCCCATCTCCATCCAGACCATGT----- 688
Db      602  CGTAGAGCTGTGGTGGGCAAGCGGATCTTCAAGCCCATCTCCATCCAGACCATGTGGGC 661

QY      689  ----- 688
Db      662  CACACTCCAGGTAATTGCACCAACATGTAGAGCAGCTCTAGGACGGGCTTGTACCGGG 721

QY      689  ----- 688
Db      722  TGGCAGTGGCCCTCACCCTGGGGCCAGCCACTACAGAGAGAGACTGACACTCGAACAGAGCTG 781

QY      689  ----- 688
Db      782  CCTCAATGAGTGAACGCGATATGGCCGACCTGAGTCTTCCGGCCTTCCAGCGCCGAGCC 841

QY      689  -----GTCCTCAGACAGAGAGCAGATGGAGCAGGGGATCGTGTGAGCTGTGGAAGT 743
Db      842  TGGCGGGTCTCTCAGAACAGGAGCAGATGGAGCAGGGGATCGTGTGAGCTGTGGAAGT 901
```

```
QY      744  GTTGGATGTCAGTAGACCTGGAGAGTGTCACTTCCAAAGAGATCCCGCAGGCTGTGGAGCT 803
Db      902  GTTGGATGTCAGTAGACCTGGAGAGTGTCACTTCCAAAGAGATCCCGCAGGCTGTGGAGCT 961

QY      804  GCGCCTGGGGCTTCCCTCCAGCAGTACCGTGTACTTCATCGACACCAATATGCTGTGCT 863
Db      962  GCGCCTGGGGCTTCCCTCCAGCAGTACCGTGTACTTCATCGACACCAATATGCTGTGCT 1021

QY      864  GGTGGCAGAGCGGAGCCGAGGCTCCCGCATCTTCCCGCAGCTTACCTGGGCTCAGAGTG 923
Db      1022  GGTGGCAGAGCGGAGCCGAGGCTCCCGCATCTTCCCGCAGCTTACCTGGGCTCAGAGTG 1081

QY      924  GAAAGCAGCAAACTTGGAGAGCTGCAGAGGAACAGGGTACCCACATCTTGAACATGCG 983
Db      1082  GAAAGCAGCAAACTTGGAGAGCTGCAGAGGAACAGGGTACCCACATCTTGAACATGCG 1141

QY      984  CCGGAGATTTGACAACTTCTACCTTGAAGGCTTCACTTACCAACAATGTGCGCTCTGGGA 1043
Db      1142  CCGGAGATTTGACAACTTCTACCTTGAAGGCTTCACTTACCAACAATGTGCGCTCTGGGA 1201

QY      1044  TGAGAGTGGGCGCCAGCTGCTGCGCACTGGAAAGAGACGACACCGCTTCAATTGAGGCTGC 1103
Db      1202  TGAGAGTGGGCGCCAGCTGCTGCGCACTGGAAAGAGACGACACCGCTTCAATTGAGGCTGC 1261

QY      1104  AAGAGCAGAGGGCACCCAGTGTCTGTCACATGCAAGATGGGCTCAGCCGCTCAGCGGC 1163
Db      1262  AAGAGCAGAGGGCACCCAGTGTCTGTCACATGCAAGATGGGCTCAGCCGCTCAGCGGC 1321

QY      1164  CACAGTGTGCGCTATGCCATGAAGCAGTTCGAAATCAGCCTTGGAGCAGGCCCTTGGGCCA 1223
Db      1322  CACAGTGTGCGCTATGCCATGAAGCAGTTCGAAATCAGCCTTGGAGCAGGCCCTTGGGCCA 1381

QY      1224  CCGTGCAGAGCTCCGGCCCATGCGCCGCCCAACCTTGGCTTCTGCGCAGCTGCAGAT 1283
Db      1382  CCGTGCAGAGCTCCGGCCCATGCGCCGCCCAACCTTGGCTTCTGCGCAGCTGCAGAT 1441

QY      1284  CTAACAGGGCATCTGACG----- 1302
Db      1442  CTAACAGGGCATCTGACG----- 1501

QY      1303  -----GCG 1305
Db      1502  GGTCTCCCGCAGAGAGACACCAGCCCTGAAAGTCTGTACACCAATTCCACTTCTTCGCGC 1561

QY      1306  AGAACCTGAGAGGTGTGGGAGAGAGAGTTGTAGGATGAGAGAGAGCAGGACAGCCCGC 1365
Db      1562  AGAACCTGAGAGGTGTGGGAGAGAGAGTTGTAGGATGAGAGAGAGCAGGACAGCCCGC 1621

QY      1366  GAAAGAAAGACCTGGGGCCACGGGGCAGCTATTAACCTCCGAGGGGTCAATGAGGTCCATC 1425
Db      1622  GAAAGAAAGACCTGGGGCCAC--GGCCACGTATTAACCTCCGAGGGGTCAATGAGGTCCATC 1679

QY      1426  AGTCTTCTGAGCGCTTCTTGGAGAGCTGGAAGAGACCTCAGTAGAGACCACTGATCCGAG 1485
Db      1680  AGTCTTCTGAGCGCTTCTTGGAGAGCTTGGAGAGAGACCTCAGTAGAGACCACTGATCCGAG 1737

QY      1486  AGGTCTTCTTCTCCAGAGTCTTTCACATGAAGAGAGCTTGGAGCCCTTCCACAGACTTG 1545
Db      1738  AGGTCTTCTTCTCCAGAGTCTTTCACATGAAGAGAGCTTGGAGCCCTTCCACAGACTTG 1797

QY      1546  CAAGGACCAAGGAGGCGCAGCAGGTGAGACAGGGGGCTCAGCCTGCGCTGAAAGTCCCGCC 1605
Db      1798  CAAGGACCAAGGAGGCGCAGCAGGTGAGACAGGGGGCTCAGCCTGCGCTGAAAGTCCCGCC 1857

QY      1606  AGTCATGTGTACCTTCCAGCGAGTCCGCTGTGGCCAAACCGGAGCCGAGGCTTCCAGG 1665
Db      1858  AGTCATGTGTACCTTCCAGCGAGTCCGCTGTGGCCAAACCGGAGCCGAGGCTTCCAGG 1917

QY      1666  AGCAGAGCAGAGGGGACAGGGGACAGGGGACAGGAGAGCCTTGCATTTTCTTACGCCAGGT 1725
Db      1918  AGCAGAGCAGAGGGGACAGGGGACAGGGGACAGGAGAGCCTTGCATTTTCTTACGCCAGGT 1977

QY      1726  TCCGGAAGGTGTGAGACAGGCCAGCGTGTATGACAGTGGAGAGAGAGGGCGAGGCGCTGAG 1785
```

```
Db 1978 TCCGGAAGTGTGAGACAGCCGACGTCATGACAGTGGAGAGAGGCGAGGCTGAG 2037
QY 1786 CCCGACACATGCCCAGGCTCCCTGACACTGAGAGATTCACAACCTCTTGGAGAAAC 1845
Db 2038 CCCGACACATGCCCAGGCTCCCTGACACTGAGAGATTCACAACCTCTTGGAGAAAC 2097
QY 1846 ACCCTCAGCTGTGTGGCGACACATCTCTCAGCTCCGGCCCATACCGCTCCTACAG 1905
Db 2098 ACCCTCAGCTGTGTGGCGACACATCTCTCAGCTCCGGCCCATACCGCTCCTACAG 2157
QY 1906 CCTCAGCTCCACCCCTGTGCTACAGGCTCCTCCTCCACCCCTGTGCTACAGCTCAG 1965
Db 2158 CCTCAGCTCCACCCCTGTGCTACAGGCTCCTCCTCCACCCCTGTGCTACAGCTCAG 2217
QY 1966 CTCTACAGCTTAAAGTCCAGGCGCCATGTCTGCTCTCCAGAGGCTCAGACTTTCTAA 2025
Db 2218 CTCTACAGCTTAAAGTCCAGGCGCCATGTCTGCTCTCCAGAGGCTCAGACTTTCTAA 2277
QY 2026 CTGGGATGTGTAGAGGAGTGAAGTACCTTTGGGGCAACAGACACCTTATTCATTC 2085
Db 2278 CTGGGATGTGTAGAGGAGTGAAGTACCTTTGGGGCAACAGACACCTTATTCATTC 2337
QY 2086 TCAACTCTAGCCCTGACACTCCTGTGCGACGGAATGAAACAGAGCTTCCGTCGAA 2145
Db 2338 TCAACTCTAGCCCTGACACTCCTGTGCGACGGAATGAAACAGAGCTTCCGTCGAA 2397
QY 2146 AAGGGTCAGGCTCCGACCCCGCCGCTCCCTGACACTCCTGTGCTCCGTCGAACTCAT 2205
Db 2398 AAGGGTCAGGCTCCGACCCCGCCGCTCCCTGACACTCCTGTGCTCCGTCGAACTCAT 2457
QY 2206 TCCCTGGAACAGCAGGCGCAGGCAACCAAGTGGCCGCCAAAGGCGAGAGATCCTCAGGC 2285
Db 2458 TCCCTGGAACAGCAGGCGCAGGCAACCAAGTGGCCGCCAAAGGCGAGAGATCCTCAGGC 2517
QY 2266 CCCAGCCGCGGAGGCTGTGAAGGCTGGCAGATCGCTTCCCTCATCCACTCCACCGGTC 2335
Db 2518 CCCAGCCGCGGAGGCTGTGAAGGCTGGCAGATCGCTTCCCTCATCCACTCCACCGGTC 2577
QY 2326 CAGGCTTTGCTGTCGCCAGACCTCTGTGACACAGGCGAGTACAGGCGACAG 2385
Db 2578 CAGGCTTTGCTGTCGCCAGACCTCTGTGACACAGGCGAGTACAGGCGACAG 2637
QY 2386 GCCAGAGATAGTCTTCTTTTGTCTTGTGCTGCTGTGCTGCTGCTGCTGCTGCTGCT 2445
Db 2638 GCCAGAGATAGTCTTCTTTTGTCTTGTGCTGCTGTGCTGCTGCTGCTGCTGCTGCT 2697
QY 2446 TACAGTATCTGCTTGTACTGAGAAATAAACACATTTTCATATTTGTTAAAAAAA 2505
Db 2698 TACAGTATCTGCTTGTACTGAGAAATAAACACATTTTCATATTTGTTAAAAAAA 2757
QY 2506 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2540
Db 2758 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2792
```

RESULT 5
US-09-955-732-1

```
; Sequence 1, Application US/09955732
; Publication No. US20020182203A1
; GENERAL INFORMATION:
; APPLICANT: Mel, Bo
; APPLICANT: Lucie, Ralf M.
; TITLE OF INVENTION: DSP-15 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.433
; CURRENT APPLICATION NUMBER: US/09/955,732
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1980
; TYPE: DNA
; ORGANISM: Homo sapiens
```

US-09-955-732-1

Query Match 53.0%; Score 1345.6; DB 9; Length 1980;
Best Local Similarity 85.1%; Pred. No. 0;
Matches 1689; Conservative 0; Mismatches 4; Indels 291; Gaps 5;

```
QY 88 ATGACCCCTGTGTCACAGTGTAGGCGTTTCGCCCCCGGCGACGCGGCTCCACGCCCTGGGG 147
Db 1 ATGACCCCTGTGTCACAGTGTAGGCGTTTCGCCCCCGGCGACGCGGCTCCACGCCCTGGGG 60
QY 148 CCTGGGACACAGCGGCTCCACGAGAGTGTGACTCCAGCGAGGCGAGAGCTTTGCGGTG 207
Db 61 CCTGGGACACAGCGGCTCCACGAGAGTGTGACTCCAGCGAGGCGAGAGCTTTGCGGTG 120
QY 208 CTCCGTTGGGGCTGTCTGAGGCTCAGAGTGAAGGGAATGATGATGCGAGAGGCGC 267
Db 121 CTCCGTTGGGGCTGTCTGAGGCTCAGAGTGAAGGGAATGATGATGCGAGAGGCGC 180
QY 268 AGTTCTGAGCCACAGAGAAAGCCCGAGTGAAGAGAGTCTCACGGGAGACAGACAGAC 327
Db 181 AGTTCTGAGCCACAGAGAAAGCCCGAGTGAAGAGAGTCTCACGGGAGACAGACAGAC 240
QY 328 TTGCGGACAGAGATCCACAGAGTCCACAGAGAGAGAGAGAGAGAGCTCAGCTC 387
Db 241 TTGCGGACAGAGATCCACAGAGTCCACAGAGAGAGAGAGAGAGAGAGCTCAGCTC 300
QY 388 ATGTTACAGCTGTGAGGCGCAGAGATGACATCCGCTGGGCGACGCTGGAGGCAACC 447
Db 301 ATGTTACAGCTGTGAGGCGCAGAGATGACATCCGCTGGGCGACGCTGGAGGCAACC 360
QY 448 CGGCTCCCGCGCTCCGCTACCTGCTGATGTTCTACACAGAGAGAGAGAGTCTGAGC 507
Db 361 CGGCTCCCGCGCTCCGCTACCTGCTGATGTTCTACACAGAGAGAGAGAGTCTGAGC 420
QY 508 CAGATGAGAGAGTCTCTGCGGCTGATTTCCCTGACAGAGCTCCAGCTGACAC 567
Db 421 CAGATGAGAGAGTCTCTGCGGCTGATTTCCCTGACAGAGCTCCAGCTGACAC 480
QY 568 CTGGGCTGGCTGTGCGGCTCTGAGTGAACCCAGAGGTACTAGATGAGAGAGGAGG 627
Db 481 CTGGGCTGGCTGTGCGGCTCTGAGTGAACCCAGAGGTACTAGATGAGAGAGGAGG 540
QY 628 TTGACGCTGACGCTGTGTTGGGCAAGCGGATCTTCAAGCCCATCTCATCCAGACCATG 687
Db 541 TTGACGCTGACGCTGTGTTGGGCAAGCGGATCTTCAAGCCCATCTCATCCAGACCATG 600
QY 688 T----- 688
Db 601 TGGGCGACACTCCAGTATTGCAACAGCATGTGAGGAGCTTAGGCGAGGCGCTTGTGA 660
QY 689 ----- 688
Db 661 CCGGGTGGAGTGCCTTCACCTGAGGCGACCTTACAGAGAGAGACTGMACTCCGAACAG 720
QY 689 ----- 688
Db 721 AGCTGCTCAATAGTGGAGCGGTATGCGGCGACTTGGAGTCTGCGGCTCCAGGCGC 780
QY 689 -----GTTCTCAGAACAGAGACAGATGAGAGAGGCGGATCGTGTGAGCTGTG 738
Db 781 GAGCCTGGCGGCTCTCAGAACAGAGACAGATGAGAGAGGCGGATCGTGTGAGCTGTG 840
QY 739 AAGTGTGATGATGACAGTGTGAGAGTGTGCTGCTCAAGAGAGATCCGCGAGGCTGTG 798
Db 841 AAGTGTGATGATGACAGTGTGAGAGTGTGCTGCTCAAGAGAGATCCGCGAGGCTGTG 900
QY 799 GAGCTGGCGCTGGGGCTCCCGCTCCAGAGTACCTGACTTCAATGAGAACAGAGATGCTG 858
Db 901 GAGCTGGCGCTGGGGCTCCCGCTCCAGAGTACCTGACTTCAATGAGAACAGAGATGCTG 960
QY 859 CTGCTGTGACACAGCGGAGCGAGCTCCGCAATCTTCCCACTCTTACCTGGGCTCA 918
Db 961 CTGCTGTGACACAGCGGAGCGAGCTCCGCAATCTTCCCACTCTTACCTGGGCTCA 1020
```


Db 781 GAGCCTGGGGGTCTCTAGAACAGAGCAGATGAGAGCGGATCCGCTGAGCTGTGG 840
QY 739 AAAGTGTGATGTCAGTACCTGGAGAGTGTCACTTCCAAAGAGATCCGCGAGCTCTG 798
Db 841 AAAGTGTGATGTCAGTACCTGGAGAGTGTCACTTCCAAAGAGATCCGCGAGCTCTG 900
QY 799 GAGCTGCGCCTGGGGCTCCCTCCAGCAGTACCTGATGATGATGATGATGATGATGATG 858
Db 901 GAGCTGCGCCTGGGGCTCCCTCCAGCAGTACCTGATGATGATGATGATGATGATGATG 960
QY 859 CTGCTGTGTGACAGAGCGGAGCCGATCCCGATCTTCCCGACCTTACCTGGGCTCA 918
Db 961 CTGCTGTGTGACAGAGCGGAGCCGATCCCGATCTTCCCGACCTTACCTGGGCTCA 1020
QY 919 GAGTGAAGCAGCAAACTGGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 978
Db 1021 GAGTGAAGCAGCAAACTGGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
QY 979 ATGCGCCGGAGATGACAACTTACCTGAGCGCTTACCTACCAATGATGAGCGCTC 1038
Db 1081 ATGCGCCGGAGATGACAACTTACCTGAGCGCTTACCTACCAATGATGAGCGCTC 1140
QY 1039 TGGGATGAGAGTGGCGCCAGCTGCTGCGCACTGGAAGAGAGAGAGAGAGAGAGAG 1098
Db 1141 TGGGATGAGAGTGGCGCCAGCTGCTGCGCACTGGAAGAGAGAGAGAGAGAGAGAG 1200
QY 1099 GCTGCAAGAGCAGAGGAGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1158
Db 1201 GCTGCAAGAGCAGAGGAGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
QY 1159 GCGGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1218
Db 1261 GCGGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
QY 1219 CGCCACGTGACAGAGCTCGGCGCCATGCGCGCCCAACCTGCTGCTGCTGCTGCTGCT 1278
Db 1321 CGCCACGTGACAGAGCTCGGCGCCATGCGCGCCCAACCTGCTGCTGCTGCTGCTGCT 1380
QY 1279 CAGATCTACGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1314
Db 1381 CAGATCTACGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1416

RESULT 7
US-09-969-708-398/C
Sequence 398, Application US/09969708
Patent No. US20020102532A1
GENERAL INFORMATION:
APPLICANT: Augustus, Meena
TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
FILE REFERENCE: 689290-70
CURRENT APPLICATION NUMBER: US/09/969, 708
PRIOR FILING DATE: 2001-10-03
PRIOR APPLICATION NUMBER: US/60/237, 606
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: US/60/237, 608
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: US/60/237, 425
NUMBER OF SEQ ID NOS: 658
SOFTWARE: Patentin version 3.0
SEQ ID NO 398
LENGTH: 334
TYPE: DNA
ORGANISM: Homosapiens
US-09-969-708-398

Query Match 12.88; Score 324.4; DB 10; Length 334;
Best Local Similarity 98.28; Pred. No. 1,7e-70;
Matches 328; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2179 TGCACCTCTGCTCTCTCTCCAGTTCATTCCTGGAACACAGCCAGGCGAGGACCAAGTGGC 2238
Db 334 TGCACCTCTGCTCTCTCTCCAGTTCATTCCTGGAACACAGCCAGGCGAGGACCAAGTGGC 275
QY 2239 CCCCCAAGGACAGAGATCTCAGAGCCCGAGCCGCGAGGCTGGAGAGGCTGGCAGAT 2298
Db 274 CCCCCAAGGACAGAGATCTCAGAGCCCGAGCCGCGAGGCTGGAGAGGCTGGCAGAT 215
QY 2299 GCGTCCCTCATTCACCTCCACCGGCTCCAGTCTTTGCTGTCTCCAGACCTCTCTG 2358
Db 214 GCGTCCCTCATTCACCTCCACCGGCTCCAGTCTTTGCTGTCTCCAGACCTCTCTG 155
QY 2359 ACACACCGCAGATCACAGGCGCAGGCGCAGAGATAGTCTTTGCTGTCTCTGCGC 2418
Db 154 ACACACCGCAGATCACAGGCGCAGGCGCAGAGATAGTCTTTGCTGTCTCTGCGC 95
QY 2419 CTGTGGCTAGTCACTTTTCATTCAGCTTACAGTATCTGCTTTGCTGAGAAATTAAC 2478
Db 94 CTGTGGCTAGTCACTTTTCATTCAGCTTACAGTATCTGCTTTGCTGAGAAATTAAC 35
QY 2479 ACATTTTCATTTGCTTAATAAAAAAAAAAAAAA 2512
Db 34 ACATTTTCATTTGCTTAATAAAAAAAAAAAAAA 1

RESULT 8
US-09-775-925-1
Sequence 1, Application US/09775925
Patent No. US20010049358A1
GENERAL INFORMATION:
APPLICANT: Lucche, Ralf M.
APPLICANT: Mei, Bo
TITLE OF INVENTION: DSP-12 AND DSP-13 DUAL-SPECIFICITY
FILE REFERENCE: 200125.420
CURRENT APPLICATION NUMBER: US/09/775, 925
PRIOR FILING DATE: 2001-02-01
NUMBER OF SEQ ID NOS: 33
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 1949
TYPE: DNA
ORGANISM: Homo sapiens
US-09-775-925-1

Query Match 8.38; Score 212; DB 10; Length 1949;
Best Local Similarity 58.98; Pred. No. 2.1e-42;
Matches 365; Conservative 0; Mismatches 255; Indels 0; Gaps 0;

QY 690 GTCTCGAAGACAGAGCAGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 749
Db 406 GCGCAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 465
QY 750 TGTGAGTACCTGAGAGTGTCACTTCCAAAGAGATCCGCGAGCTGTGAGCTGCGCT 809
Db 466 GAGCAGAGATGTAAATAATGTACTTCCAAAGAGATTCGTAAATGTAGAGAAACAGAT 525
QY 810 GGGGCTCCCTCCAGCAGTACCTGACTTATGACACACAGATGCTGTGTGTGGC 869
Db 526 GAATTTGTAATTTGAAGGAAGTCAAGGAATTTATAGCAATGATGTACTTATCTTGGG 585
QY 870 ACAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 929
Db 586 ACAGATGAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 645
QY 930 AGCAACCTGAGAGAGTGTGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 989
Db 646 ATCCATCTGAGAGAGTGTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 705
QY 990 GATTGACACTTACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1049
Db 706 AATGATTAATTTTCTGCTGCTTATTTGATATCATATACATTCGAGATCATGATGAGAG 765

[illegible]

```

RESULT 9
; US-09-775-925-7
; Sequence 7, Application US/09775925
; Patent No. US20010049358A1
; GENERAL INFORMATION:
; APPLICANT: Lucche, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-12 AND DSP-13 DUAL-SPECIFICITY
; TITLE OF INVENTION: PHOSPHATASES
; FILE REFERENCE: 200125.420
; CURRENT APPLICATION NUMBER: US/09/775,925
; CURRENT FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1052
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-775-925-7

```

Query Match	8.28;	Score 208.4;	DB 10;	Length 1052;
Best Local Similarity	57.18;	Pred. No. 1.2e-41;		
Matches 380; Conservative	0;	Mismatches 286;	Indels 0;	Gaps 0;

[illegible]

QY	1115	GCACCCAGTGTCTGTGTCCACTGCAAGATGGGCGCTACGCCGCTACAGGGGCACAGTGTCTGG	1174
Db	620	GATCTAAATGCTTGTGTCTACTGCAAAATGGGGGTGAGTGCTGTACGCTGTACACCGGATTTG	679
QY	1175	CCATATGCCATGAACACGTACGAAATGCAGCCGTGGAGCAGGCCCTTGCGCCACGTGCAGAGAC	1234
Db	680	CCATATGCAATGAAGGAAATATGGCTGGAAATCTGGACCGAGCTATGACTGTGAAGAAAT	739
QY	1235	TCCGGCCATGTGCCCCCGCCCAACCTGTGGCTCTGTGGCGCCAGCTGCAGATCTACCGAGGCA	1294
Db	740	GACGAAAGGTAAACCAAGCCCAACCCAAAGCTTCATGAGAACACTGGAAAGATGTCAGGGCA	799
QY	1295	TCTCTGACCGCAGCAACCTGAGGGGTGTGGGGAGAAAGATTGTATGGCATGTGAACAAGAGC	1354
Db	800	TCTTGTGTGGCAAGCTTCTTAGGGCTTGATTCTATGGAGGAGGGGCAAAAGCCCTTGGGAGAGA	859
QY	1355	CAGGCA 1360	
Db	860	AAAGCA 865	

```

RESULT 10
; US-09-775-925-5
; Sequence 5, Application US/09775925
; Patent No. US20010049358A1
; GENERAL INFORMATION:
; APPLICANT: Lucche, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-12 AND DSP-13 DUAL-SPECIFICITY
; TITLE OF INVENTION: PHOSPHATASES
; FILE REFERENCE: 200125.420
; CURRENT APPLICATION NUMBER: US/09/775,925
; CURRENT FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1711
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-775-925-5

```

Query Match	8.2%	Score 208.4	DB 10	Length 1711
Best Local Similarity	57.1%	Pred. NO. 1.5e-41		
Matches 380; Conservative	0	Mismatches 286	Indels 0	Gaps 0

Qy	695	CAAAAGGAGCAGATGGAGAGGCGATCCGCTGCTGAGCTGTGGAAAGTGTTGGATGTCA	754
Db	859	CTGAACGTGAACGAACAGAAAGCTAATTAAACCAAAATTAAGGGGATCATGATGCAGA	918
Qy	755	GTCAGCTTGGAGAGTGTCATCTTCCAAAGATATCCGCAAGGCTGTGGAGTGCAGCCCTGGGGC	814
Db	919	AGGATTTGGAAATATTACATCTCAAGAGATTAAAGAACAGAGTTGGAAATGCAAATGTGTGT	978
Qy	815	TCGCCCTCCAGACAGTACCCTGACTTCATGCACAACAGATGCTGCTGTGTGGCACAAGC	874
Db	979	GCAACTTCGGGGAATTCAGAGAAATTTATAGCAATGAATAGATGATCTTTGGCTCAA	1038
Qy	875	GGGAGCCAGGAGCTCCCCGATCTTCCCCACTCTACTGGGCTCAGAGTGGAAACGCAGCAA	934
Db	1039	TGGATAGGCTTACACAGATATTGAGAGATGTGTTCTCGGGCTCAGATGGAAATGGCTCCA	1098
Qy	935	ACCTGAGGAGCTCTCAGAGGAGCAAGGGTCAACCAATCTGGAACATAGGGCCGGAGATTG	994
Db	1099	ACTTAGAGAGACTTACAGAACCGAGGGGTACGGTATATCTTTGAATGTCACTCCAGAGATAG	1155
Qy	995	ACAACCTTCTACCTGAGAGCGCTTCACTTACCACCATGTGCGCTCTGGGAGTGAAGGATCGG	1054
Db	1159	ATAAATCTTCTCCAGAGAGTCTTTGATATCATATAAATTTGGGGATATATGATGAAGAGCAA	1218
Qy	1055	CCCACTCTGCGCCCACTCGAGAGAGACGACCGCTTCATTGAGGCTGCAGAGACACAGG	1114
Db	1219	CGGATCTCTGCGCTACTGAGATATACCTTAACAAATATCATCTCTAAGACAAAGAACTGT	1278

QY 1115 GCACCCAGCTGCTGCTCCTGCAAGATGGGCGTCAAGCCGCTCAGCGGCCACAGTCTGG 1174
 Db 1279 GATCTAATATGCTTGTGACACTGCAAAATGGGGGTGAGTGGCTCAGCCCTCCAGCCTGATTG 1338
 QY 1175 CCTATGCTGATGAGACATGACATGCAAGCCTGGAGCAGAGCCCTCGCCAGCTGCAAGAGC 1234
 Db 1339 CCTATGCAATGAAAGATGATGCTGGAATCTGGACCGAGCCTATGACTATGCTGAAAGAAA 1398
 QY 1235 TCCGGCCATGCGCCCGCCCAACCTGGGCTTCTGCGCCAGCTCAGATCTACAGAGGCA 1294
 Db 1399 GACGAAAGGTATACCAAGCCCAAGCCTTTCATGAGACATCGAAGAGTATAGGGGA 1458
 QY 1295 TCCGTAGCGCCCAAGCTGAGGCTGGTGGGAGAGAAAGTTAGGCAATGGAAGAGAGC 1354
 Db 1459 TCTTGCTGCAAGCTTCTAGGCTTGTATTCATGAGAGGAGGACAAAGCCCTGGGAGAGA 1518
 QY 1355 CAGGCA 1360
 Db 1519 AAGCA 1524

RESULT 11

US-10-002-600-51
 ; Sequence 51, Application US/10002600
 ; Patent No. US20020137077A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hopkins, Christopher M.
 ; APPLICANT: Peterson, David P.
 ; APPLICANT: Cocks, Benjamin G.
 ; APPLICANT: Hawkins, Phillip R.
 ; TITLE OF INVENTION: GENES REGULATED IN ACTIVATED T CELLS
 ; FILE REFERENCE: PA-0042 US
 ; CURRENT APPLICATION NUMBER: US/10/002,600
 ; PRIOR FILING DATE: 2001-10-25
 ; PRIOR APPLICATION NUMBER: 60/243,521
 ; NUMBER OF SEQ ID NOS: 116
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 51
 ; LENGTH: 1685
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; OTHER INFORMATION: Template ID: 235631.2
 US-10-002-600-51

Query Match 3.9%; Score 98.8; DB 12; Length 1685;

Best Local Similarity 53.3%; Pred. No. 1.6e-14;
 Matches 208; Conservative 0; Mismatches 182; Indels 0; Gaps 0;

QY 892 ATCTTCCCACTCTACCTGAGGCTCAGAGTGAAGCAGCAACCTGAGAGAGTGCAG 951
 Db 613 ATCTTCCCACTCTACCTGAGGCTCAGAGTGAAGCAGCAACCTGAGAGAGTGCAG 672
 QY 952 AGAAGAGAGGCTCAGCAATCTTGAACATGAGCGCCGAGATTGCAACTTCTACCTGAG 1011
 Db 673 GCCGTGAGCATCAAGCGCTCTCAAGCTGTCCCGCAGCTGCCCACTTGGAGGCG 732
 QY 1012 CGCTTCACTACCAATGTGCGCTCTGGAGTGAAGTCCGCCAGCTGCTCCGCGAC 1071
 Db 733 CTCTTCCGCTCAAGATATCTCTGGAGTGAAGCAACAGATGTGAGATCAGTCCGCG 792
 QY 1072 TGAAGAGAGCGCAGCTTCAATGAGGCTGCAAGAGCAGCAGCGCCAGCCAGTCTGCTG 1131
 Db 793 TTTCCAGGAGGCGCATAGCTTATTTGACTGGGTGAAGAAACAGCGCGGCTGCTGGTG 852
 QY 1132 CACTGCAAGATGAGGCTGAGCGGCTCAGCGGCAAGTGTGCTGCTATGCAATGAAGCAG 1191
 Db 853 CACTGCAAGGAGGCTGAGCGGCTCAGCGGCAAGTGTGCTGCTATGCAATGAAGCAGT 912
 QY 1192 TACGAATGACCTGAGAGAGCGGCTGCGCGCAGTGAAGAGATCCGCGCCAGTCCGCGC 1251

Db 913 CGCGGTGCGGCTGAGACAGGCGCTTGTACTTCTTAAGACGCGCGGGGCTCATCTCC 972
 QY 1252 CCGAAGCCTGGCTCTCTGCGCCAGCTGCGAG 1281
 Db 973 CCGAAGCCTGGCTCTCTGCGCCAGCTGCGAGTCTGCTG 1002

RESULT 12

US-09-808-701-12
 ; Sequence 12, Application US/09808701
 ; Patent No. US20020146757A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Zhou, Ping
 ; APPLICANT: Goodrich, Ryle
 ; APPLICANT: Asundi, Vinod
 ; APPLICANT: Zhao, Qing A.
 ; APPLICANT: Wang, Jian-Rui
 ; APPLICANT: Xue, Aidong J.
 ; APPLICANT: Ren, Feiyan
 ; APPLICANT: Wang, Dunrui
 ; APPLICANT: Chen, Rui-hong
 ; APPLICANT: Dmanac, Radoje T.
 ; TITLE OF INVENTION: No. US20020146757A1 Nucleic Acids and
 ; FILE REFERENCE: 790CIP2D
 ; CURRENT APPLICATION NUMBER: US/09/808,701
 ; PRIOR FILING DATE: 2001-03-14
 ; PRIOR APPLICATION NUMBER: 09/649,167
 ; PRIOR FILING DATE: 2000-08-23
 ; PRIOR APPLICATION NUMBER: 09/540,217
 ; NUMBER OF SEQ ID NOS: 17
 ; SOFTWARE: PL_FL_genes Version 2.0
 ; SEQ ID NO 12
 ; LENGTH: 969
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (34)..(765)
 US-09-808-701-12

Query Match 3.8%; Score 96; DB 10; Length 969;

Best Local Similarity 56.4%; Pred. No. 6.2e-14;
 Matches 202; Conservative 0; Mismatches 150; Indels 6; Gaps 1;

QY 349 CCCCAGAGAGGAGGAGCAGAGCAGACCTGACCTCATGTATACAGCTGTGAGGCGG 408
 Db 235 CACAAGCATGAGGTGATCTGCTCAACATCTTCAGTGTATGATCAACCTTCTGCGTTGC 294
 QY 409 CAGATGACATCCGCTGAGCGCCAGCTGAGAGCAGACCCGCGCTCCCGGCTCCGCTAC 468
 Db 295 GAAAGCAAGAACTTCAAGCTGAGCGCTGAGAGCGCGCTGGGGGAGCGGGTCCGGTAC 354
 QY 469 CTGCTGTAGTGTCTTACACGAGAGGAGAAAGTGTGAGCCAGATGAGAGCGTCTCTCTG 528
 Db 355 ATGCTGTGTGTGTACAGCAGCGGCGCGCAGAG-----CACGAGGAGAAATATCTTCTG 408
 QY 529 GCGGTGATTTTCCCTGACAGCAGCTCCCGCAGCTGCACACCTGTGGCTGTGCTCCCTC 588
 Db 409 GGAATGACCTTTTCAGTAAGGAAAGTAAGAGTGCACCATTTGGAGTGTCTCCGACTG 468
 QY 589 TGAAGTACACCCAGGTGTACTTGTATGAGAGCGGGGCTTCAAGCTGTGAGTGTGG 648
 Db 469 TGAAGTACACCCAGGTGTACTTGTATGAGAGCGGGGCTTCAAGCTGTGAGTGTGG 528
 QY 649 CAAAGCGGATCTTCAAGCCCATCTCCATCCAGACATGTGTCTCTAGAGACAGAGAGC 706
 Db 529 AGATGACATATTTTAAGCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 586

RESULT 13

```
US-09-865-993-3
; Sequence 3, Application US/09865993
; Publication No. US20030060437A1
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: Andrew T. Watt
; TITLE OF INVENTION: ANTISENSE MODULATION OF DUAL SPECIFIC PHOSPHATASE 5 EXPRESSION
; FILE REFERENCE: RTS-0175
; CURRENT APPLICATION NUMBER: US/09/865,993
; CURRENT FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 89
; SEQ ID NO 3
; LENGTH: 2139
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (47)...(1201)
US-09-865-993-3
```

```
Query Match          3.5%; Score 89.2; DB 9; Length 2139;
Best Local Similarity 51.5%; Pred. No. 4.3e-12;
Matches 205; Conservative 0; Mismatches 193; Indels 0; Gaps 0;
```

```
QY 892 ATCTTCCCCCACTCTTACCTGGGCTCAGAGTGAAGCAAGCAAACTGGAGAGCTGCAG 951
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 590 ATCTTCCCCCTCTCTACCTTGAAGTGCCTACCATGCATCCAGTGCAGATTCCTCGCC 649

QY 952 AGGAACAGGTCACCCACATCTTGAAACATGGCCCGGAGATTGACACTTCTACCTTGAG 1011
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 650 AACTTGACATCAGACCCCTGTAATGTCTCCGACGAGCTCCGAGGCTCGAATGAC 709

QY 1012 CGCTTCACTACACACAAATGTGGGCTCTGGATGAGAGTGGCCACCTGCTGCCGAC 1071
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 710 CACCTACACTACAAATGTAGTCCCTTGAAGACACACACAGCGCTGCATTTAGTCCAC 769

QY 1072 TGAAGAGAGAGCAGCCCTTATTAGGCTGCAGAGACAGCAGGCACTGCTGCTG 1131
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 770 TTTCAAGAGCAATAGACTTCTATGTGCTGTGACAGGAAAGGAGGCAAGGCTCTGCTG 829

QY 1132 CACTGCAGATGGGGCTGAGCGGCTCAGCGGCAACAGTGGCTATGCGCATGAAGCAG 1191
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 830 CACTGTGAGGCTGGGATCTCCCGTTACCCACCATCTGCACTGCTTATTTGAAGAC 889

QY 1192 TAGCAATGAGCAGCTGAGCAGGCGCTGCGCCACGTGAGAGAGCTCGGCCCATCGCCGC 1251
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 890 AAGCAGTTCCGCTGAAGAGGCTTTCATATCATCAATCAAGAGAGAGCATGTGCTCG 949

QY 1252 CCCAACCCTGGCTTCTGCGCCAGCTGCAGATCTACCA 1289
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 950 CCCAACTTTGGCTTCATGTGGCGCAGCTCTGCAGTACGA 987
```

```
RESULT 14
; Sequence 10, Application US/09865993
; Publication No. US20030060437A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Brett P. Monia
; APPLICANT: Andrew T. Watt
; TITLE OF INVENTION: ANTISENSE MODULATION OF DUAL SPECIFIC PHOSPHATASE 5 EXPRESSION
; FILE REFERENCE: RTS-0175
; CURRENT APPLICATION NUMBER: US/09/865,993
; CURRENT FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 89
; SEQ ID NO 10
; LENGTH: 2479
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
```

```
US-09-865-993-10
```

```
Query Match          3.5%; Score 89.2; DB 9; Length 2479;
```

```
Best Local Similarity 51.5%; Pred. No. 4.6e-12;
Matches 205; Conservative 0; Mismatches 193; Indels 0; Gaps 0;
```

```
QY 892 ATCTTCCCCCACTCTTACCTGGGCTCAGAGTGAAGCAAGCAAACTGGAGAGCTGCAG 951
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 758 ATCTTCCCCCTCTCTACCTTGAAGTGCCTACCATGCATCCAGTGCAGATTCCTCGCC 817

QY 952 AGGAACAGGTCACCCACATCTTGAAACATGGCCCGGAGATTGACAACTTCTACCTTGAG 1011
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 818 AACTTGACATCAGACCCCTGTAATGTCTCCGACGAGCTCCGAGGCTCGATGAC 877

QY 1012 CGCTTCACTACACAAATGTGGGCTCTGGATGAGAGTGGCCACCTGCTGCCGAC 1071
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 878 CACCTACACTACAAATGTAGTCCCTTGAAGACACACACAGCGCTGCATTTAGTCCAC 937

QY 1072 TGAAGAGAGAGCAGCCCTTATTAGGCTGCAGAGACAGCAGGCACTGCTGCTG 1131
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 938 TTTCAAGAGCAATAGACTTCTATGTGCTGTGACAGGAAAGGAGAGGCAAGGCTCTGCTG 997

QY 1132 CACTGCAGATGGGGCTGAGCGGCTCAGCGGCAACAGTGGCTATGCGCATGAAGCAG 1191
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 998 CACTGTGAGGCTGGGATCTCCCGTTACCCACCATCTGCACTGCTTATTTGAAGAC 1057

QY 1192 TAGCAATGAGCAGCTGAGCAGGCGCTGCGCCACGTGAGAGAGCTCGGCCCATCGCCGC 1251
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1058 AAGCAGTTCCGCTGAAGAGGCTTTCATATCATCAAGAGAGAGAGCATGTGCTCG 1117

QY 1252 CCCAACCCTGGCTTCTGCGCCAGCTGCAGATCTACCA 1289
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1118 CCCAACTTTGGCTTCATGTGGCGCAGCTCTGCAGTACGA 1155
```

```
RESULT 15
US-09-902-941-1918
; Sequence 1918, Application US/09902941
; Patent No. US20020172952A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tonglong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Rietter, Marc W.
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Carter, Darick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvik, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
```

```
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.478C17
; CURRENT APPLICATION NUMBER: US/09/902,941
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 2002
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1918
; LENGTH: 1209
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-902-941-1918
```

```
Query Match          3.5%; Score 88; DB 9; Length 1209;
Best Local Similarity 50.5%; Pred. No. 6.5e-12;
Matches 214; Conservative 0; Mismatches 210; Indels 0; Gaps 0;
```

```
QY 892 ATCTTCCCCCACTCTTACCTGGGCTCAGAGTGAAGCAAGCAAACTGGAGAGCTGCAG 951
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 616 ATCTTCCCCCTCTCTACCTGCGCAGTGCCTACCATGCTCGCCGAGAGCATGCTCGAC 675

QY 952 AGGAACAGGTCACCCACATCTTGAAACATGGCCCGGAGATTGACAACTTCTACCTTGAG 1011
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 676 GCCCTGGCATCAGGCTCTGTGTAATGTCTCTCGAGACTGCCCAACCACTTTGAAGGA 735
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 10, 2003, 23:58:55 ; Search time 6901.21 Seconds
(without alignments)
5960.776 Million cell updates/sec

Title: US-09-761-640-3
Perfect score: 2540
Sequence: 1 cctgcctcctgcggctccagc.....aaaaaaaaaaaaaaaaaaaaa 2540

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estcov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	851.4	33.5	1004	BM910493
2	795	31.3	1024	BM560192
3	753.8	29.7	772	BQ014446
4	718.2	28.3	1131	BE737389
5	662.8	26.1	943	BM563401
6	659.4	26.0	991	BM471207

7	657	25.9	790	12	BG326709	BG326709	602425578
8	654	25.7	663	12	BG753072	BG753072	602732341
9	642	25.3	653	12	BG679981	BG679981	602626694
10	631	24.8	716	12	BE907514	BE907514	601497473
11	622.8	24.5	932	12	BG340220	BG340220	602438408
12	597.8	23.5	1228	14	BM921421	BM921421	601675476
13	593.2	23.4	890	12	BE901875	BE901875	601675476
14	586.4	23.1	589	9	AL602806	AL602806	DKF2P866L
15	569	22.4	672	12	BG831015	BG831015	602766736
16	562.6	22.1	791	12	BE881145	BE881145	601492016
17	559	22.0	1125	12	BM809903	BM809903	601158526
18	558.6	22.0	599	10	BE280203	BE280203	601158526
19	558	22.0	678	13	BM047611	BM047611	603628949
20	550	21.7	622	12	BE898514	BE898514	601681423
21	548.4	21.6	584	13	BM018531	BM018531	603646365
22	548	21.6	892	12	BF971565	BF971565	602239811
23	537.2	21.1	1061	14	BM906413	BM906413	602239811
24	526.8	20.7	728	13	BI824278	BI824278	603040675
25	517	20.4	937	14	BQ718968	BQ718968	602850749
26	516	20.3	719	10	BE543087	BE543087	601069024
27	508.4	20.0	917	14	BO712114	BO712114	ACENICOURT
28	498.8	19.6	863	12	BG749256	BG749256	602708085
29	492.8	19.4	496	10	BE348287	BE348287	hw21108.x
30	490.4	19.3	969	10	BE563116	BE563116	601335734
31	482.4	19.0	568	13	BI087083	BI087083	602850749
32	479	18.9	505	13	BM083928	BM083928	imageqc.2
33	477	18.8	574	13	BI007055	BI007055	OV3-RT006
34	473.2	18.6	535	9	AA134926	AA134926	z023d02.S
35	472.4	18.6	846	13	BI820326	BI820326	603036896
36	467.8	18.4	893	12	BG253488	BG253488	602364213
37	464.6	18.3	750	10	BE531347	BE531347	601278540
38	460.8	18.1	860	12	BF127735	BF127735	601810422
39	459.4	18.1	560	13	BI007133	BI007133	OV3-RT006
40	459	18.1	459	9	AI763368	AI763368	*155904.X
41	456.6	18.0	612	13	BM017088	BM017088	603643638
42	455	17.9	962	12	BG166835	BG166835	602345005
43	453	17.8	897	12	BE873337	BE873337	601450290
44	450.8	17.7	790	13	BM008005	BM008005	603617745
45	449	17.7	729	13	BI086720	BI086720	602850066

ALIGNMENTS

RESULT 1
LOCUS BM910493 1004 bp mRNA linear EST 12-MAR-2002
DEFINITION AGENCOURT_6703004 NIH_MGC_98 Homo sapiens cDNA clone IMAGE:5453997
ACCESSION BM910493
VERSION BM910493.1 GI:19360872
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1004)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
Plate: LNCMI948 row: 9 column: 22
High quality sequence stop: 725.
Location/Qualifiers
1..1004

FEATURES
source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5453997"
/tissue.lib="NIH_MGC_98"
/lab_host="astrocytoma grade IV, cell line"
/Note="Organ: brain; Vector: pORF7; Site:1: XhoI; Site:2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

```

BASE COUNT 211 a 299 c 324 g 167 t 3 others

Query Match 33.5%; Score 851.4; DB 14; Length 1004;
Best Local Similarity 97.4%; Pred. No. 3.7e-122;
Matches 875; Conservative 0; Mismatches 22; Indels 1; Gaps 1;

```

OY 159 GCGGCTCCAGCAGAGGAGTGCATCCAGCAGAGGAGCAGACTTGGCGTCTCCGCGGGC 218
DB 1 GCGGCTCCAGCAGAGGAGTGCATCCAGCAGAGGAGCAGACTTGGCGTCTCCGCGGGC 218
OY 219 TGTCTGGGACTCCAGGATGAGAGGGACATGATGATGAGAGAGGCGCAGTTGAGCC 278
DB 61 TGTCTGGGACTCCAGGATGAGAGGGACATGATGATGAGAGAGGCGCAGTTGAGCC 278
OY 279 AACAGAGAGAGGCGCCGAGTGAAGAGAGTCCAGCGGAGCAGACAGACTTGGCGGCAAG 338
DB 121 AACAGAGAGAGGCGCCGAGTGAAGAGAGTCCAGCGGAGCAGACAGACTTGGCGGCAAG 338
OY 339 ATCCAGAGATCCCGCAGAGAGAGAGAGAGAGCAGACAGTCCAGTCTGATGACGT 398
DB 181 ATCCAGAGATCCCGCAGAGAGAGAGAGAGAGCAGACAGTCCAGTCTGATGACGT 398
OY 399 GCTGAGGCGCAGAGTGCATCCGCTGGCAGCCAGCTTGAAGCAACCCGCGCTCCCG 458
DB 241 GCTGAGGCGCAGAGTGCATCCGCTGGCAGCCAGCTTGAAGCAACCCGCGCTCCCG 458
OY 459 GCTCGGCTACCTGCTGATGTTCTACAGAGAGAGAGAGTCTGAGCAGAGATGAGAC 518
DB 301 GCTCGGCTACCTGCTGATGTTCTACAGAGAGAGAGAGTCTGAGCAGAGATGAGAC 518
OY 519 GGTCTCTCTGGGCGTGAATTTCCCTGACAGAGACTCCCGAGCTCCAGCTGGCGTGT 578
DB 361 GGTCTCTCTGGGCGTGAATTTCCCTGACAGAGACTCCCGAGCTCCAGCTGGCGTGT 578
OY 579 CTGCGGCTCTGAGTGAACACCCAGGTGACTTAAATGATGAGAGAGGCGCTCAGCGTAC 638
DB 421 CTGCGGCTCTGAGTGAACACCCAGGTGACTTAAATGATGAGAGAGGCGCTCAGCGTAC 638
OY 639 GTCTGTGGGCAAGCGGATCTTCAAGCCCATCTCCATCCAGACCATGTGGTCTGAGA 698
DB 481 GTCTGTGGGCAAGCGGATCTTCAAGCCCATCTCCATCCAGACCATGTGGTCTGAGA 698
OY 699 ACAGAGAGAGATGAGAGAGGCGGATCGTGTGAGCTGTGAAATGTTGATGATGAGTGA 758
DB 541 ACAGAGAGAGATGAGAGAGGCGGATCGTGTGAGCTGTGAAATGTTGATGATGAGTGA 758
OY 759 CCTGAGAGTGTCACTTCCAAAGAGATCCGAGGCTGTGAGACTCGGCTGGGGCTCC 600
DB 601 CCTGAGAGTGTCACTTCCAAAGAGATCCGAGGCTGTGAGACTCGGCTGGGGCTCC 600
OY 819 CCTCAGAGATGAGAGTGTCACTTCCAAAGAGATCCGAGGCTGTGAGACTCGGCTGGGGCTCC 660
DB 661 CCTCAGAGATGAGAGTGTCACTTCCAAAGAGATCCGAGGCTGTGAGACTCGGCTGGGGCTCC 660
OY 879 CCGAGCTTCCCGGATCTTCCCGACCTTACCTGAGGCTCAGAGTGAAGCGACGAAACCT 938
DB 721 CCGAGCTTCCCGGATCTTCCCGACCTTACCTGAGGCTCAGAGTGAAGCGACGAAACCT 938

```

```

OY 939 GGAGAGCTGCAGAGAGACAGGCTACCCACATTTGAACAT-GGCCCGGAGATTGACA 997
DB 781 GGAGAGCTGCAGAGAGAGACAGGCTACCCACATTTGAACATGGCGGGGAGATTGACA 840
OY 998 ACTTGTACCTTCCGAGGCTTCACTTACCAATGAGGCGCTTGGATGAGAGTGGC 1055
DB 841 ACTTGTACCTTCCGAGGCTTCACTTACCAATGAGGCGCTTGGATGAGAGTGGC 898

```

RESULT 2
BM560192
LOCUS BM560192
DEFINITION BM560192 1024 bp mRNA linear EST 20-FEB-2002
ACCESSION AGENOURL_6597753 NIH_MGC_98 Homo sapiens cDNA clone IMAGE:5431968
VERSION BM560192
KEYWORDS BM560192.1 GI:18804414
SOURCE EST.
ORGANISM human.
Eukaryote: Homo sapiens

REFERENCE
AUTHORS Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
1 (bases 1 to 1024)
TITLE NIH-MGC http://mgc.ncl.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: ATCC

CDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LNLN)
Clone distribution by: Agencourt Bioscience Corporation
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.lnl.gov
Plate: LCM1905 row: b column: 01
high quality sequence stop: 655.

FEATURES
source location/Qualifiers
1..1024

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5431968"
/tissue.lib="NIH_MGC_98"
/lab_host="astrocytoma grade IV, cell line"
/Note="Organ: brain; Vector: pORF7; Site:1: XhoI; Site:2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

```

BASE COUNT 221 a 305 c 328 g 170 t

Query Match 31.3%; Score 795; DB 13; Length 1024;
Best Local Similarity 98.4%; Pred. No. 1.9e-113;
Matches 824; Conservative 0; Mismatches 10; Indels 3; Gaps 2;

```

OY 158 AGCGGCTCCAGCAGAGGAGTGCATCCAGCAGAGGAGCAGACTTGGCGTCTCCGCGGG 217
DB 1 AGCGGCTCCAGCAGAGGAGTGCATCCAGCAGAGGAGCAGACTTGGCGTCTCCGCGGG 217
OY 218 CTGTCTGGGAGTGCAGATGAGAGGGACATGATGATGAGAGAGGCGCAGTTGAGC 277
DB 61 CTGTCTGGGAGTGCAGATGAGAGGGACATGATGATGAGAGAGGCGCAGTTGAGC 277
OY 278 CAACAGAGAGGCGCCGAGTGAAGAGAGTCCAGCGGAGCAGACAGACTTGGCGGCAAG 337
DB 121 CAACAGAGAGGCGCCGAGTGAAGAGAGTCCAGCGGAGCAGACAGACTTGGCGGCAAG 337
OY 338 GATCCAGAGTCCCGCAGAGAGAGAGAGAGAGTCCAGCTTACCTGATGATGATGAGC 397
DB 121 GATCCAGAGTCCCGCAGAGAGAGAGAGAGAGTCCAGCTTACCTGATGATGATGAGC 397

```


QY 2402 TTTTGTCTTCTGCGCTCTGCTAGTCAATTTTTCATAGCCTTACAGTATCTGCTTT 2461
|||||
Db 113 TTTTGTCTTCTGCGCTCTGCTAGTCAATTTTTCATAGCCTTACAGTATCTGCTTT 54
|||||
QY 2462 GTACTGGAATAAACAACATTTTCATATTGGTTAAAAA 2514
|||||
Db 53 GTACTGGAATAAACAACATTTTCATATTGGTTAAAAA 1
|||||
RESULT 4
BE737389 1131 bp mRNA linear EST 15-SEP-2000
LOCUS BE737389
DEFINITION mRNA sequence.
ACCESSION BE737389
VERSION BE737389.1 GI:10151381
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1131)
NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM344 row: h column: 03
High quality sequence stop: 795.
Location/Qualifiers
1..1131
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3640370"
/clone_lib="NIH-MGC-39"
/tissue_type="adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: pancreas; Vector: pCMV7; Site:1: XhoI;
Site:2: EcoRI; CDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed
by Ling Hong in the laboratory of Gerald M. Rubin
(University of California, Berkeley) using ZAP-CDNA
synthesis kit (Stratagene) and Superscript II RT (Life
Technologies)."
BASE COUNT 298 a 373 c 271 g 189 t
ORIGIN
Query Match 28.3%; Score 718.2; DB 12; Length 1131;
Best Local Similarity 94.4%; Pired. No. 1.3e-101;
Matches 777; Conservative 0; Mismatches 43; Indels 3; Gaps 3;
QY 1684 GGCAGGGGAGGAGGAGCCCTGATTTCTCTACGCCAGCTTCGGAAGAGTGTGAGAC 1743
|||||
Db 5 GGGCGGGGAGTGAAGAGCTGATTTCTCTACGCCAGCTTCGGAAGAGTGTGAGAC 64
|||||
QY 1744 AGGCACAGCGATGACATGAGAGAGGAGGCGCTGAGCCCTACACATGCCACAG 1803
|||||
Db 65 AGGCACAGCGATGACATGAGAGAGGAGGCGCTGAGCCCTACACATGCCACAG 124
|||||
QY 1804 CTCCCTGACACTGAGAGAGATCCACAACTCTTGAGAAAACCCCTACAGTCTGTTGCC 1863
|||||
Db 125 CTCCCTGACACTGAGAGAGATCCACAACTCTTGAGAAAACCCCTACAGTCTGTTGCC 184
|||||
QY 1864 GCACACATTCCTCTCAGCTCGGCCCATACCCGCTACTACAGCTCACTCCACCCCTG 1923
|||||

Db 185 GCACACATTCCTCTCAGCTCGGCCCATACCCGCTACTACACACCTCCACCTCCG 244
|||||
QY 1924 TCACATAGGCGCTCAGCTCCAGCCCGCTGACATACAGCCCTCCTCAGCTTAAGTC 1983
|||||
Db 245 TCACATAGGCGCTCAGCTCCAGCCCGCTGACATACAGCCCTCCTCAGCTTAAGTC 304
|||||
QY 1984 CCAGGCCCATGTCGCTGCTGCTCCAGAGGCTCAAGACTTTTCACTGGATGTGTAGAGG 2043
|||||
Db 305 CCAGGCCCATGTCGCTGCTGCTCCAGAGGCTCAAGACTTTTCACTGGATGTGTAGAGG 364
|||||
QY 2044 ACTGAAGTACCTTTGGGGGCAACAGACCCATGTTTCACTTCACTTACCTCTGAC 2103
|||||
Db 365 ACTGAAGTACCTTTGGGGGCAACAGACCCATGTTTCACTTCACTTACCTCTGAC 424
|||||
QY 2104 ACTCAGCTGAGAGGATGAAGAAACAGAGCTTCCCGTCAAAAAGGTCAGCCCTCCA 2163
|||||
Db 425 ACTCAGCTGAGAGGATGAAGAAACAGAGCTTCCCGTCAAAAAGGTCAGCCCTCCA 484
|||||
QY 2164 CCCCCCGCCCTCCCTGACCTCTGCTCTCCAGTTCACTTCTGGAACCCAGCCAGG 2223
|||||
Db 485 CCCCCCGCCCTCCCTGACCTCTGCTCTCCAGTTCACTTCTGGAACCCAGCCAGG 544
|||||
QY 2224 CAGGCACAGAGTGGCCCCCAAGGAGGAGATCTCAGGCGCCAGCGGGAGGCTG 2283
|||||
Db 545 CAGGCACAGAGTGGCCCCCAAGGAGGAGATCTCAGGCGCCAGCGGGAGGCTG 604
|||||
QY 2284 GAAGGCTGGAGAGTGGCTTCCCTATCCAGCTCCAGCGGTCAGTCTTCTGCTGTC 2343
|||||
Db 605 GAAGGCTGGAGAGTGGCTTCCCTATCCAGCTCCAGCGGTCAGTCTTCTGCTGTC 663
|||||
QY 2344 CCCAGACTCTCTGTGACACCGACCGAGATCAGAGGACCCAGGAGAGATGCTCTT 2403
|||||
Db 664 CCAGAG-CTCTGTGTGACACCGACCGAGATCAGAGGAC-CAAGCCAGAGATGCTCTT 721
|||||
QY 2404 TTTGCTCTTCTGCGCTCTGCTGCTAGTCAATTTTTCATAGCCTTACAGTATCTGCTTT 2463
|||||
Db 722 TTTGCTCTTCTGCGCTCTGCTGCTAGTCAATTTTTCATAGCCTTACAGTATCTGCTTT 781
|||||
QY 2464 ACTGAGAAATAAACAACATTTTCATATTGGTTAAAAA 2506
|||||
Db 782 ACTGAGAAATAAACAACATTTTCATATTGGTTAAAAA 824
|||||
RESULT 5
BM563401 943 bp mRNA linear EST 20-FEB-2002
LOCUS BM563401
DEFINITION AGENCOURT_6589387 NIH_MGC_98 Homo sapiens CDNA clone IMAGE:5441645
5', mRNA sequence.
ACCESSION BM563401
VERSION BM563401.1 GI:18810292
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 943)
NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1916 row: e column: 06
High quality sequence stop: 618.
Location/Qualifiers
1..943
FEATURES
source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5441645"
/clone_lib="NIH_MGC_98"
/tissue_type="astrocytoma grade IV, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: brain; Vector: pOTR1; Site: 1: XhoI; Site: 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAC(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library."

BASE COUNT 201 a 284 c 305 g 153 t
ORIGIN

Query Match 26.1%; Score 662.8; DB 13; Length 943;
Best Local Similarity 94.2%; Pred. No. 5e-93;
Matches 710; Conservative 0; Mismatches 42; Indels 2; Gaps 2;

OY 159 GGGGCTCCAGGAGAGATCCACTCCAGCAAGAGACCTTGGCGTCCCTGGGGC 218
DB 26 GGGGCTCCAGGAGAGATCCACTCCAGCAAGAGACCTTGGCGTCCCTGGGGC 85
OY 219 TGTCTGGAGCTGAGATGAGAGGAGACATGATGTCAGACAGAGCGACGTTCTGAGCC 278
DB 86 TGTCTGGAGCTGAGATGAGAGGAGACATGATGTCAGACAGAGCGACGTTCTGAGCC 145
OY 279 AACAGAGAAGGCCCCAGTGAAGAGAGAGCTCCACGGGAGCAGACAGACTTGGGCAAG 338
DB 146 AACAGAGAAGGCCCCAGTGAAGAGAGAGCTCCACGGGAGCAGACAGACTTGGGCAAG 205
OY 339 ATCCAGAGTCCCGAAGCAGAGAGAGAGACAGACAGCTGCACTCATGTGTCACGT 398
DB 206 ATCCAGAGTCCCGAAGCAGAGAGAGAGACAGACAGCTGCACTCATGTGTCACGT 265
OY 399 GGTGAGGCGCAGATGATCCGCTGGAGCCAGCTGAGAGGACCCGGCCTCCCG 458
DB 266 GGTGAGGCGCAGATGATCCGCTGGAGCCAGCTGAGAGGACCCGGCCTCCCG 325
OY 459 GCTCCGCTACTGCTGTGATGTTTCTACAGAGAGAGAGAGGTCGAGCCAGATGAGAC 518
DB 326 GCTCCGCTACTGCTGTGATGTTTCTACAGAGAGAGAGAGGTCGAGCCAGATGAGAC 385
OY 519 GGTCTCTGCGGCGGTGATTCCTCGACAGAGAGCTCCCGACCTGAGCCCTGGT 578
DB 386 GGTCTCTGCGGCGGTGATTCCTCGACAGAGAGCTCCCGACCTGAGCCCTGGT 445
OY 579 CTTCGCTCTGAGATGACACCAAGGTACTTAGATGAGAGCGGGGCTTCAGCGTGAC 638
DB 446 CTTCGCTCTGAGATGACACCAAGGTACTTAGATGAGAGCGGGGCTTCAGCGTGAC 505
OY 639 GTTGTGTGGGAAAGCCGATCTTCAGGCCATCTCATCCAGACATGTGGTCTCAGA 698
DB 506 GTTGTGTGGGAAAGCCGATCTTCAGGCCATCTCATCCAGACATGTGGTCTCAGA 565
OY 699 ACAGAGAGATGAGAGAGAGAGAGATCCGCTGAGAGAGAGAGAGAGAGAGAGAGAG 758
DB 566 ACAGAGAGATGAGAGAGAGAGAGATCCGCTGAGAGAGAGAGAGAGAGAGAGAGAG 625
OY 759 CTTGAGAGATGATCTTCAAGAGATCCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 817
DB 626 CTTGAGAGATGATCTTCAAGAGATCCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 685
OY 818 CCTTCACAGAGATAC - GTGACTTCATGACAAACAGAGATGCTGCTGGTGACAGAGG 876
DB 686 CCTTCACAGAGATACCGGTGACTTCATGACAAACAGAGATGCTGCTGGTGACAGAG 745
OY 877 GAGCAGAGCTCCCGATCTTCCCTCCACCTTACC 910
DB 746 GCGGGGAGCCGAGCCCTTCCTCCGATCTTTC 779

RESULT 6
LOCUS BM471207 991 bp mRNA linear EST 05-FEB-2002
DEFINITION AGENCOURT_6478155 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5563111
ACCESSION BM471207
VERSION BM471207.1 GI:18520249
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cga@bbs-remail.nih.gov
Tissue Procurement: ATCC/DCPD/DRP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LLM12293 row: j column: 08
High quality sequence stop: 592.

FEATURES
source
1..991
Location/Qualifiers

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5563111"
/clone_lib="NIH_MGC_72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pCMV-SPORT6; Site: 1: NotI;
Site: 2: SalI; Cloned unidirectionally. Primer: oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."

BASE COUNT 199 a 303 c 329 g 159 t 1 others
ORIGIN

Query Match 26.0%; Score 659.4; DB 13; Length 991;
Best Local Similarity 99.0%; Pred. No. 1.6e-92;
Matches 663; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 34 GTTGAAGGAAGGGGCGTCCCGGTGACAGCCAGCTGCTGGGCTCATGAGCC 93
DB 1 GTTGAAGGAAGGGGCGTCCCGGTGACAGCCAGCTGCTGGGCTCATGAGCC 60
OY 94 CTGTGACAGTGAAGCGGTTCCCGCCGAGCGGCGCTTCAGCGCCGCTGG 153
DB 61 CTGTGACAGTGAAGCGGTTCCCGCCGAGCGGCGCTTCAGCGCCGCTGG 120
OY 154 GAGCAGGCGGTCCAGAGAGAGTGAATCCAGAGAGAGAGAGAGAGAGAGAGAG 213
DB 121 GAGCAGGCGGTCCAGAGAGAGTGAATCCAGAGAGAGAGAGAGAGAGAGAGAG 180
OY 214 GGGGCTGTCTGGGACTCCAGAGATGAGAGGAGCAATGATGATGACAGAGAGAGAG 273
DB 181 GGGGCTGTCTGGGACTCCAGAGATGAGAGGAGCAATGATGATGACAGAGAGAGAG 240
OY 274 GAGCAG 333
DB 241 GAGCAG 300
OY 334 CAAGGATCCAGAGAGTCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 393
DB 301 CAAGGATCCAGAGAGTCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
OY 394 CAGTGTGAGAGCGCAGAGATGATCCGCTGGAGCCAGCTGAGAGAGAGAGAGAGAG 453

Db 361 CAGCTGCTGAGGCGGCGAGATGACATCCGCTGGACGCCAGCTGGAGGACCCCGGCT 420
QY 454 CCGCGGCTCCGCTACCTGCTGTAGTTTCTACAGAGAGAGAGAGTGTAGCCAGAT 513
Db 421 CCGCGGCTCCGCTACCTGCTGTAGTTTCTACAGAGAGAGAGTGTAGCCAGAT 480
QY 514 GAGAGGCTCTCTGGGCGTGAGATTTCCCTGACAGCAGCTCCCGACCTGACCCGCGG 573
Db 481 GAGAGGCTCTCTGGGCGTGAGATTTCCCTGACAGCAGCTCCCGACCTGACCCGCGG 540
QY 574 CTGGCTCTGGGCGGCTGGAGAGACACCCAGGTACTTATATGAGAGCGGGGCTTACG 633
Db 541 CTGGCTCTGGGCGGCTGGAGAGACACCCAGGTACTTATATGAGAGCGGGGCTTACG 600
QY 634 GTGACGCTGTGGTGGGCAAGCCGAGTCTTCAGCCCATCTTCATCCAGACATGTGTCC 693
Db 601 GTGACGCTGTGGTGGGCAAGCCGAGTCTTCAGCCCATCTTCATCCAGACATGTGTCC 660
QY 694 TCAGACAGG 703
Db 661 ACACTCCAGG 670

RESULT 7
BG326709 790 bp mRNA linear EST 27-FEB-2001
LOCUS 602425578F1 NIH_MGC_14 Homo sapiens cDNA clone IMAGE:4563595 5',
ACCESSION mRNA sequence.
VERSION BG326709
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 790)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTF

cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LNCMI277 row: k column: 20
High quality sequence stop: 737.

FEATURES
source Location/Qualifiers

1..790
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="4563595"
/clone_lib="NIH_MGC_14"
/tissue_type="renal cell adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: kidney; Vector: pORF7; site:1: XhoI; site:2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(5). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 162 a 237 c 268 g 123 t

ORIGIN
Query Match 25.9%; Score 657; DB 12; Length 790;
Best Local Similarity 99.2%; Pred. No. 4.2e-92;
Matches 660; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 39 GGGAGAGGGGCGCTGCCCGGTGCGACGCCAGGTGCTGCGGGCTGCTCCATGAGCCCTGGT 98
Db 1 GGGAGAGGGGCGGCGGCGGTGCGACGCCAGGTGCTGCGGGCTGCTCCATGAGCCCTGGT 60
QY 99 CACAGTGAACCGGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 158
Db 61 CACAGTGAACCGGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
QY 159 GCGGTCACAGCAGAGTGCAGCTCCAGCGAAGGCGAGACTTTGGGTCGCTCGGCGGCG 218
Db 121 GCGGTCACAGCAGAGTGCAGCTCCAGCGAAGGCGAGACTTTGGGTCGCTCGGCGGCG 180
QY 219 TGTCTGGAGCTGCAGATGAGAGGGAGCAATGATGATGATGATGATGATGATGATGAT 278
Db 181 TGTCTGGAGCTGCAGATGAGAGGGAGCAATGATGATGATGATGATGATGATGATGATGAT 240
QY 279 AACAGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 338
Db 241 AACAGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
QY 339 ATCCAGAGTCCCGCAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 398
Db 301 ATCCAGAGTCCCGCAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
QY 399 GCTGAGGCGCGCAGAGATGACATCCGCTGGCAGCCAGCTGAGAGGACCCCGGCTCCCG 458
Db 361 GCTGAGGCGCGCAGAGATGACATCCGCTGGCAGCCAGCTGAGAGGACCCCGGCTCCCG 420
QY 459 GCTCCGCTACCTGCTGT 518
Db 421 GCTCCGCTACCTGCTGT 480
QY 519 GGTCTCTCTGGGCGGTGATTTCCCTGACAGCAGCTCCCGACCTGACCTGGGCTGGT 578
Db 481 GGTCTCTCTGGGCGGTGATTTCCCTGACAGCAGCTCCCGACCTGACCTGGGCTGGT 540
QY 579 CTTGCCCTCTGTGAGTGACACCCAGGTGTACTAGATGAGAGCGGGGCTTCAAGCGTAC 638
Db 541 CTTGCCCTCTGTGAGTGACACCCAGGTGTACTAGATGAGAGCGGGGCTTCAAGCGTAC 600
QY 639 GTTGTGTGGGCAAGCGGAGTCTTCAAGCCCATCTTCATCCAGACCATGTGTCTCTGAG 698
Db 601 GTTGTGTGGGCAAGCGGAGTCTTCAAGCCCATCTTCATCCAGACCATGTGTGTCTCTGAG 660
QY 699 ACAGG 703
Db 661 CCAAG 665

RESULT 8
BG753072 663 bp mRNA linear EST 15-MAY-2001
LOCUS 602732241F1 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:4875562 5',
DEFINITION mRNA sequence.
ACCESSION BG753072
VERSION BG753072.1 GI:14063725
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 663)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov
Plate: LLCM1757 row: b column: 11
High quality sequence stop: 663.

BASE COUNT	124 a	205 c	231 g	103 t
ORIGIN				

Query Match	25.7%	Score 654;	DB 12;	length 663;
Best Local Similarity	100.0%	Pred. No. 1.3e-91;		
Matches 654;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	27	CCGCGGGGTTGAGGAAGGGGCGTCGGCCGGGTGCCAGCCAGAGTCTGCGGGCTTGGCTC	86
Db	10	CCGCGGGGTTGAGGAAGGGGCGTCGGCCGGGTGCCAGCCAGAGTCTGCGGGCTTGGCTC	69
QY	87	CATGGCCCTGTGCACAGTAGAGCCGTTTGGCCCCGGGCGAGCGGGCGCTCCACGCCCGTGGG	146
Db	70	CATGGCCCTGTGCACAGTAGAGCCGTTTGGCCCCGGGCGAGCGGGCGCTCCACGCCCGTGGG	129
QY	147	GCCCTGGGACCAAGGGGGTCCAGCGAAGAGTGCATCCAGCGCAAGCGAGCTTGGCGGT	206
Db	130	GCCCTGGGACCAAGGGGGTCCAGCGAAGAGTGCATCCAGCGCAAGCGAGCTTGGCGGT	189
QY	207	GCTCCGTGGGGCTGTCTCTGGGACTCAGAGATGAGAGGGGACAAATGATGATGACAGCAGAGCG	266
Db	190	GCTCCGTGGGGCTGTCTCTGGGACTCAGAGATGAGAGGGGACAAATGATGATGACAGCAGAGCG	249
QY	267	CAGTTCTGAGCCACACAGAGAGGGCCCCGAGTGTAGAGAGAGCTTCCAGGGGGACAGACAGA	326
Db	250	CAGTTCTGAGCCACACAGAGAGGGCCCCGAGTGTAGAGAGAGCTTCCAGGGGGACAGACAGA	309
QY	327	CTTGGGGCAAGGATCCACAGATGCCACAGAGGAGGAGGACAGAGGCGACACTGCACCT	386
Db	310	CTTGGGGCAAGGATCCACAGATGCCACAGAGGAGGAGGACAGAGGCGACACTGCACCT	369
QY	387	CATGTGTACAGCTGCTGAGAGGCGCAGAGATGACATTCGCGCTGGCAGGCCACAGCTGGAGGCACC	446
Db	370	CATGTGTACAGCTGCTGAGAGGCGCAGAGATGACATTCGCGCTGGCAGGCCACAGCTGGAGGCACC	429
QY	447	CCGGCCCTCCCGGCGCTCCGCTACTCTGTAATTCTTACAGCAGAGAGGAAGTCTGAG	506
Db	430	CCGGCCCTCCCGGCGCTCCGCTACTCTGTAATTCTTACAGCAGAGAGGAAGTCTGAG	489
QY	507	CCAGATGAGAGAGGTCACCTCGGCGGTGGATTTCCCTGTACACACAGCTCCCGCAGGTGCAC	566
Db	490	CCAGATGAGAGAGGTCACCTCGGCGGTGGATTTCCCTGTACACACAGCTCCCGCAGGTGCAC	549
QY	567	CCTGGGCGTGTGCTTGGCCCTCTGGAGTGCACCCAGAGTGTACTTAGATGAGAGCGGGG	626
Db	550	CCTGGGCGTGTGCTTGGCCCTCTGGAGTGCACCCAGAGTGTACTTAGATGAGAGCGGGG	609
QY	627	CTTACAGCGTACGCTGTGGTGGCAAGCCGGATCTTCAAGCCCATCTCATATCCA	680
Db	610	CTTACAGCGTACGCTGTGGTGGCAAGCCGGATCTTCAAGCCCATCTCATATCCA	663

LOCUS	653 bp	linear	EST 01-MAY-2001
DEFINITION	BG679981		
	602626694F1	NCI_CGAP_Sk4	Homo sapiens CDNA clone IMAGE:4751758 '5'
	mRNA sequence.		

ACCESSION	BG679981	
VERSION	BG679981.1	GI:13911378
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 653)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.

FEATURES	Location/Qualifiers
source	1. .653

```

/db_xref="taxon:9606"
/clone="IMAGE:4751758"
/clone_lib="NCI-CGAP_Skn4"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/notes="Organ: skin; Vector: pCMV-SPORT6; Site: 1; NotI;
Site: 2; SalI; Cloned unidirectionally. Primer: oligo dT.
Average insert size 1.5Kb. Library constructed by Life
Technologies. Note: this is a NCI-CGAP Library."

BASE COUNT      143 a      173 g      114 t
ORIGIN

```

Oy	1503	GAGTTCACATGAAGACCTCTGCAGCCCTTCCACAGCTTGCAGAGGACCAAGGAGAGC	1562
Db	1	GAGTTCACATGAAGACCTCTGCAGCCCTTCCACAGCTTGCAGAGGACCAAGGAGAGC	60
Oy	1563	CAGCAGGTGAGAGGGGGCCTCAGCCCTCCTGAAAGTCCCGCCACTGATGTTACCTTC	1622
Db	61	CAGCAGGTGAGAGGGGGCCTCAGCCCTCCTGAAAGTCCCGCCACTGATGTTACCTTC	120
Oy	1623	CAGGGCACTGCCGTGTGTGGCCCAACCGGACCAGGCTTCCAGAGACAGAGGAGGAG	1682
Db	121	CA-GGCAGTGGCGTGTGTGGCCCAACCGGACCAGGCTTCCAGAGACAGAGGAGGAG	179
Oy	1683	GGGCGAGGGGCAAGGAGAGACCCCTGCATTTCTCTACAGCCAGTTTCCGGAAGGTGTGAGA	1742
Db	180	GGGCGAGGGGCAAGGAGAGACCCCTGCATTTCTCTACAGCCAGTTTCCGGAAGGTGTGAGA	239
Oy	1743	CAGGCCACGTCATGACAGTGGAGAGGAGGCGAGGCTGAGCCCTCACACATGCCCCAC	1802
Db	240	CAGGCCACGTCATGACAGTGGAGAGGAGGCGAGGCTGAGCCCTCACACATGCCCCAC	299
Oy	1803	GCTCCCTTGACACTGAAGAAGATCCACAATCTCTTGGAGAAACACCTCAGCTCTGTTC	1862
Db	300	GCTCCCTTGACACTGAAGAAGATCCACAATCTCTTGGAGAAACACCTCAGCTCTGTTC	359
Oy	1863	CGCACACATTCCTTCACACTCGGAGCCCAATACCGGACATACAGGCTTCAGTCCGACCCCT	1922
Db	360	CGCACACATTCCTTCACACTCGGAGCCCAATACCGGACATACAGGCTTCAGTCCGACCCCT	419
Oy	1923	GTCACATGAGGGCTACGTCGCCACCCCTTGATCAGGCTTCACGCTCAACGCTTAAAGT	1982

```

Db 420 GTGACTAGAGGCTACCTCCACCCCTGTCTACTACAGCTTCTTACAGCTTAAGT 479
QY 1983 CCAGAGCCCATGTCTGCTGTGCCAAGGCTCAGACTTTTAACTGGATGTGTAGAG 2042
Db 480 CCGAGGCCCATGTCTGCTGTGCCAAGGCTCAGACTTTTAACTGGATGTGTAGAG 539
QY 2043 GACTGAAGGTAACCTTTTGGGGGCAACAGCACCCTAGTTTCTTCAACTTACCTTGA 2102
Db 540 GACTGAAGGTAACCTTTTGGGGGCAACAGCACCCTAGTTTCTTCAACTTACCTTGA 599
QY 2103 CACTCACCCTGTGGCAGGATGAAGAGAGCTTCCCTGCAAAAAGGCTCAG 2156
Db 600 CACTCACCCTGTGGCAGGATGAAGAGAGCTTCCCTGCAAAAAGGCTCAG 653

```

```

RESULT 10
BE907514 716 bp mRNA linear EST 20-OCT-2000
LOCUS 601497473F1 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:3899418 5',
DEFINITION mRNA sequence.
ACCESSION BE907514
VERSION BE907514.1 GI:10401153
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

```

```

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC

```

```

CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM9697 row: a column: 19

```

```

High quality sequence stop: 683.
Location/Qualifiers
1..716

```

```

FEATURES
source

```

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NIH_MGC_70"
/clone_image="IMAGE:3899418"
/issue_type="epithelioid carcinoma"
/lab_host="PH10B (phage-resistant)"
/notes="Organ: pancreas; Vector: pCMV-SPORT6; Site: 1; NotI;
Site: 2; SalI; Cloned unidirectionally. Primer: oligo dr.
Average insert size 1.1 kb. Library constructed by Life
Technologies."
BASE COUNT 138 a 213 c 251 g 114 t
ORIGIN

```

```

Query Match 24.8% Score 631; DB 12; Length 716;
Best Local Similarity 96.6%; Pred. No. 4.5e-88;
Matches 687; Conservative 0; Mismatches 20; Indels 4; Gaps 4;

```

```

QY 11 CCGGTCCAGAGACTGCTCCGCGGGTTAGAGGAAGGCGCGTGGCCAGTGGCAGCCAGGT 70
Db 1 CCGGTCCAGAGACTGCTCCGCGGGTTAGAGGAAGGCGCGTGGCCAGTGGCAGCCAGGT 60
QY 71 GCTCCGCGCTGGCTGCATGGCCCTGTGTACAGTGAACCGTTCGCCCGGAGAGCGCG 130
Db 61 GCTCCGCGCTGGCTGCATGGCCCTGTGTACAGTGAACCGTTCGCCCGGAGAGCGCG 119
QY 131 CCTCAGACCCCTGGGCGCTGGAGACAGCGGTCTCAGCAAGAGGTCGATCCAGCGAA 190
Db 120 CCTCAGACCCCTGGGCGCTGGAGACAGCGGTCTCAGCAAGAGGTCGATCCAGCGAA 179

```

```

QY 191 GGCAGAGCTTTGGGTGCTCCGTGGGGCTGTCTGGGACTGACAGGATGAGGGGACATG 250
Db 180 GGCAGAGCTTTGGGTGCTCCGTGGGGCTGTCTGGGACTGACAGGATGAGGGGACATG 239
QY 251 ATGATGCGAGAGAGCCAGTTCTTACGCCAAGAGAAAGGCCCGAGTGAAGAGAGCTCC 310
Db 240 ATGATGCGAGAGAGCCAGTTCTTACGCCAAGAGAAAGGCCCGAGTGAAGAGAGCTCC 299
QY 311 ACGGGAGACACAGAGACTTGGGGCAAGGATCCAGAGTCCCGAAGACAGAGAGACAGA 370
Db 300 ACGGGAGACACAGAGACTTGGGGCAAGGATCCAGAGTCCCGAAGAGAGAGACAGA 359
QY 371 GGCAGACCTGTCACTTCATGCTACAGTGTGAGCCGAGATGACATCCGCTGGCAG 430
Db 360 GGCAGACCTGTCACTTCATGCTACAGTGTGAGCCGAGATGACATCCGCTGGCAG 419
QY 431 CCCAGCTGAGAGCACCCTGGGCTCCCGGCTCGGCTACCTGCTAGTCTTCTACAGAG 490
Db 420 CCCAGCTGAGAGCACCCTGGGCTCCCGGCTCGGCTACCTGCTAGTCTTCTACAGAG 479
QY 491 AAGGAGAGGCTGAGACCCAGATGAGAGCGTCTCTGGGGCTGATTTCCCTGACAGCA 550
Db 480 AAGGAGAGGCTGAGACCCAGATGAGAGCGTCTCTGGGGCTGATTTCCCTGACAGCA 539
QY 551 GCTCCCGCAGCTGACACCCTGGGCTGTCTTGGCCCTCTGGAGTGACACCAGGTACT 610
Db 540 GCTCCCGCAGCTGACACCCTGGGCTGTCTTGGCCCTCTGGAGTGACACCAGGTACT 599
QY 611 TAGATGAGAGACGGGGGCTTCAAGCGTGTGAGTGTGGGCAAGCCGATCTCAAGCCA 670
Db 600 TAGATGAGAGACGGGGGCTTCAAGCGTGTGAGTGTGGGCAAGCCGATCTCTCAAG 657
QY 671 TCTCATCCAGACATGCTGTCTCAGACAGCAGCAGCAGGAGGAGGAGAGCCCA 721
Db 658 TCTCATCCAGACATGCTGTCTCAGACAGCAGCAGGATGTGCGCAAGCATGTGA 707

```

```

RESULT 11
BG340220 932 bp mRNA linear EST 27-FEB-2001
LOCUS BG340220
DEFINITION 602438408F1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4556401 5',
ACCESSION BG340220
VERSION BG340220.1 GI:13146647
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

```

```

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC

```

```

CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM1258 row: p column: 02
High quality sequence stop: 630.

```

FEATURES

```
source
```

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="IMAGE:4556401"
/clone_lib="NIH_MGC_46"
/issue_type="leiomyosarcoma cell line"
/lab_host="PH10B (phage-resistant)"

```

```

/notice:organ: uenrus: Vector: pOTB7. Site.1: XhoI; Site.2:
EcoRI; cDNA: made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCGCGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-CNA synthesis kit (stratagene) and SuperscriptII
II RT (Life Technologies). Note: this is a NIH-MGC
library."

```

BASE COUNT	160 a	284 c	326 g	162 t
ORIGIN				

Query Match	24.5%;	Score 622.8;	DB 12;	Length 932;
Best Local Similarity	98.6%;	Pred. No. 7.5e-87;		
Matches 649;	Conservative 0;	Mismatches 7;	Indels 2;	Gaps 2

QY	46	GGGTGCGCCGGGTGCACAGCCAGGTGCTCGGGGCTCGGCTGCATAGGCCCTGTGCACAGTG	105
Db	1	GGCCGTGCCCCGGTCCACACCAGGTGTCTCGGGCTCGGCTGCATAGGCCCTGTGCACAGTG	60
QY	106	AGCCGTTGCCCCGGGGAGCGGGCCCTCCAGCGCCCTGTGGGGCCCTTGGAGCAAGCGGTC	165
Db	61	AGCCGTTGCCCCGTGGGAGCGGGGCTCCAGCGCCGTGGGGCCCTTGGAGCAAGCGGTC	120
QY	166	CAGCGAAGGAATCCACTTCACGCCAAGGCAGAGCTTTTCGGTGCTCGGTGGGGCTGTCTTG	225
Db	121	CAGGGAAGGATCTCACTCCACGGAAGGCAGAGCTTTTCGGTGCTCGGTGGGGCTGTCTTG	180
QY	226	GGACGTGCGAGTGTGAGGGGGACAAATGATGATGACACAGGCGAGTTGTGAGCCAAACAG	285
Db	181	GGACTGCGAGTGTGAGGGGGACAAATGATGATGACACAGGCGAGTTGTGAGCCAAACAG	240
QY	286	AAGGCCCGGAGTGAAGGAGAGCTCCACGGGGACACGACAGACTTCGGGCAAGGATCCAG	345
Db	241	AAGGCCCGGAGTGAAGGAGAGCTCCACGGGGACACGACAGACTTCGGGCAAGGATCCAG	300
QY	346	AGTCCCGGAGGACGAGAGGACAGAGCAGCAGCTGCACCTCATAGGTATACAGCTGTAGG	405
Db	301	AGTCCCGGAGGACGAGAGGACAGAGCAGCAGCTGCACCTCATAGGTATACAGCTGTAGG	360
QY	406	CCGCGAGATGACATCCGCTTGGCAGCCAGCTGTGAGGCGACCCCGGCTCCCGGCTTCGCG	465
Db	361	CCGCGAGATGACATCCGCTTGGCAGCCAGCTGTGAGGCGACCCCGGCTCCCGGCTTCGCG	419
QY	466	TACCTGCTGTGATGTTTCTACACGAGAAAGAGAGGTTGAGCCAGATGAGAGCGTCTC	525
Db	420	TACCTGCTGTGATGTTTCTACACGAGAAAGAGAGGTTGAGCCAGATGAGAGCGTCTC	479
QY	526	CTGGGCGTGAATTCCTCGACAGAGAGTCCCGGCGTGCACCTCGGGGCGTGTCTTGCCC	585
Db	480	CTGGGCGTGAATTCCTCGACAGAGAGTCCCGGCGTGCACCTCGGGGCGTGTCTTGCCC	539
QY	586	CTCTGGAGTGAACACCAGAGGTGTACTTATGATGAGACCGGGGCTTCACGCTGAGCTGTGT	645
Db	540	CTCTGGAGTGAACACCAGAGGTGTACTTATGATGAGACCGGGGCTTCACGCTGAGCTGTGT	599
QY	646	GGGCAAAAGCCGGAATCTTCAAGCCCATCTCAATCCAGACCATGTGTGCTCTAGAACAG	703
Db	600	GGGCAAAAGCCGGAATCTTCAAA-CCATCTTCATCAAGACCAATGTGTGGGCGACACTCCAG	656

RESULT	12
BM921421	
LOCUS	1228 bp
DEFINITION	mRNA linear EST 12-MAR-2007
	AGENCOURT_6626274 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5752855
	5' , mRNA sequence.

REFERENCE	1 (bases 1 to 1228)
AUTHORS	NIH-MGC http://mgc.ncbi.nlm.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC).
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D.

Source

1.1228

```
/organism="Homo sapiens"  
/db_xref="taxon:9606"  
/clone="IMAGE:5752858"  
/clone_lib="NIH_MGC_115"  
/lab_host="DH10B"
```

/note="Organ: pooled brain, lung, testis; Vector: pcwv-sport6; Site:1: Notr; Site2: Ecovr (destroyed); RNA source anonymous.pool of 6 male brains, age range 23-27; male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (Ecovr site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH-MGC Library."
 259 a 390 c 345 g 231 t 3 others

BASE COUNT	259 a	390 c	345 g	231 t	3 others
ORIGIN					

Query Match	23.5%	Score 597.8;	DB 14;	Length 1228;
Best Local Similarity	94.7%	Pred. No. 4.9e-83;		
Matches 641;	Conservative 0;	Mismatches 32;	Indels 4;	Gaps 2;

QY	14	GTGAGACTGTCCCGCGGGGTGAGGGAAAGAGGGCCGTGCCGGTGCACAGCCAGAGTGT	73
Db	50	GGCCCGGAATTCCTCGGAGTGTGTTAGGGAAAGAGGGCCGGTCCAGCCACAGTGTCT	109
QY	74	CGCGGCTGTGCTCATATGAGCCCTGTGTCACATGAGAGCCCTTCCGCCCGGAGCGCCGCT	133
Db	110	CGCGGCTGTGCTCATATGAGCCCTGTGTCACATGAGAGCCCTTCCGCCCGGAGCGCCGCT	165
QY	134	CCAGCCCGCTGGGGCCCTTGGAGACCAGCGGGTCCAGCCAGAGAGTTCGACTCCAGGAAGCC	199
Db	170	CCAGCCCGCTGGGGCCCTTGGAGACCAGCGGGTCCAGCCAGAGAGTTCGACTCCAGCGGAAGCC	223
QY	194	AGAGCTTTCGGGTGCTCGGTGGGGCTGTCTCTGGGACTCAGAGATGAGAGGGACATATG	255
Db	230	AGAGCTTTCGGGTGCTCGGTGGGGCTGTCTCTGGGACTCAGAGATGAGAGGGACATATG	287
QY	254	ATGCAGCAGAGGCCATGTTCTGAGCCACAGAGAAAGGCCCGCAGTGAGAGAGGAGCTCCAG	313
Db	290	ATGCAGCAGAGGCCATGTTCTGAGCCACAGAGAAAGGCCCGCAGTGAGAGAGGAGCTCCAG	345
QY	314	GGGACCAACAGACTTCGGGGCAAGGATCCAGAGTCCCAAGCCCAACAGAGAGAGCAGAGGC	377
Db	350	GTGACCAACAGACTTCGGGGCAAGGATCCAGAGTCCCAAGCCCAACAGAGAGAGCAGAGGC	409
QY	374	AGCACCTTGACACTCATGTTAGTACAGCTGTCTGAGGCCGACAGATGACATCCGCTGTGCACCC	433
Db	410	AGCACCTTGACACTCATGTTAGTACAGCTGTCTGAGGCCGACAGATGACATCCGCTGTGCACCC	465
QY	434	AGCTGTGAGGACACCCCGGCTCCCGGGCTCCGCTACACCTGCTGTGTGTTCTTACACGAGAG	499
Db	470	AGCTGTGAGGACACCCCGGCTCCCGGGCTCCGCTACACCTGCTGTGTGTTCTTACACGAGAG	523
QY	494	GAGAAAGTCTGAGCCAGGATAGACAGGTCTCTCTGTGGGTGTGTGATTTCCCTGTACAGCAGCT	555

|||||
 Db 530 GAGAAAGTGTGAGCCAGAGATGAGAGAGGTCCTCGGCGCGTGAGATTCCCGAGACAGAGCT 589
 QY 554 CCCCCAGCTGACCCCTGGGCGCTGTCTGCCCTGTGAGATGACACCCAGCTACTAG 613
 Db 590 CCCCCAGCTGACCCCTGGGCGCTGTCTGCCCTGTGAGATGACACCCAGGCTACTAG 649
 QY 614 ATGAGACGAGGCGCTTCAGAGCT--GACGTCGTGGTGCAAA--GCCGATCTTCAAGCCC 669
 Db 650 ATGAGACGAGGCGCTTCAGAGCTGTGAGGCGGCAAAACCCGATCTCTAAGGCC 709
 QY 670 ATTCATCCAGACCAT 686
 Db 710 ATCCCTCTTAAAAAAT 726

RESULT 13

BE901875 890 bp mRNA linear EST 29-SEP-2000
 LOCUS 601675476F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3958314 5',
 DEFINITION mRNA sequence.

ACCESSION BE901875
 VERSION BE901875.1 GI:10391490

KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 890)
 NIH-MGC. <http://mgc.ncl.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-rc@mail.nih.gov
 Tissue Procurement: ATCC

CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
 Plate: LLC837 row: 9 column: 19
 High quality sequence stop: 736.
 Location/Qualifiers
 1..890

FEATURES
 source

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3958314"
 /clone_lib="NIH_MGC_21"
 /tissue_type="choriocarcinoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: Placenta; Vector: pOTB7; Site: 1: XhoI;
 Site: 2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACGAG(G). Site-selected by
 for average insert size 1.8kb. Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."
 206 a 265 c 274 g 145 t

BASE COUNT

Query Match 23.4%; Score 593.2; DB 12; Length 890;
 Best Local Similarity 95.1%; Pred. No. 2.8e-82;

Matches 711; Conservative 0; Mismatches 23; Indels 14; Gaps 9;

QY 923 GGAACGAGCAAACTGGAGAGCTGACAGAGAAAGGTCACCCACATCTTGAACATGG 982
 Db 1 GGAACGAGCAAACTGGAGAGCTGACAGAGAAAGGTCACCCACATCTTGAACATGG 59
 QY 983 CCGGGAGATTGACACTTACCCGAGGCGTTCACCTCCCAATGTGCGCTCTGGG 1042
 Db 60 CCGGGAGATTGACACTTACCCGAGGCGTTCACCTCCCAATGTGCGCTCTGGG 119

QY 1043 ATGAGAGTGTGCGCCAGCTGCTGCCGCACTGGAAGAGAGACGACCCGCTTCATTAGAGCTG 1102
 Db 120 ATGAGAGTGTGCGCCAGCTGCTGCCGCACTGGAAGAGAGACGACCCGCTTCATTAGAGCTG 179
 QY 1103 CAAGAGCACAGGCGACCCACAGCTGCTGCTGCTGCAAGATGGCGTTCAGCGG 1162
 Db 180 CAAGAGCACAGGCGACCCACAGCTGCTGCTGCTGCAAGATGGCGTTCAGCGG 239
 QY 1163 CCACAGTGTGCGCTTCATGCAATGACAGTACGATGACGCTGGAGCAGAGCCCTGCGG 1222
 Db 240 CCACAGTGTGCGCTTCATGCAATGACAGTACGATGACGCTGGAGCAGAGCCCTGCGG 299
 QY 1223 ACGTGACAGAGCTCGGCGCCATGCGCCGCCAACCCTGCTTCTGCGCCAGCTGACAG 1282
 Db 300 ACGTGACAGAGCTCGGCGCCATGCGCCGCCAACCCTGCTTCTGCGCCAGCTGACAG 359
 QY 1283 TCTACCAAGGCGATCTGAGGCGCCAGAACTGAGGTTGTGGGAGAGAAAGTTGTAGGC 1342
 Db 360 TCTACCAAGGCGATCTGAGGCGCCAGAACTGAGGTTGTGGGAGAGAAAGTTGTAGGC 418
 QY 1343 ATGGAAGAGAGCGGCGAGCGCCGAAAGAGAGCTGGGCGCCAGGCGACGTATAACC 1402
 Db 419 ATGGAAGAGAGCGGCGAGCGCCGAAAGAGAGCTGGGCGCCAGGCGACGTATAACC 475
 QY 1403 TCCGAGGCGTCATGAGGTCATCATCTTCTGAGAGCCCTCTTGGAGGCTGAGAGCACCC 1462
 Db 476 TCCGAGGCGTCATGAGGTCATCATCTTCTGAGAGCCCTCTTGGAGGCTGAGAGCACCC 534
 QY 1463 TCAGTAGACCACTGACATGCGCAGAGGTTCTTCTTCCACAGGTTTCAATGAAAGGCC 1522
 Db 535 TCAG-TGAGCAGAGCATGATCCAGAGGCTCTTCTTCCACAGGTTTCAATGAAAGGCC 593
 QY 1523 TCTGAGGCGCTTCCAGAGCTTCCAGAGGCAAGGAGGCGACAGGAGGAGGCGG 1582
 Db 594 TCTGAGGCGCTTCCAGAGCTTCCAGAGGCAAGGAGGCGACAGGAGGAGGCGG 652
 QY 1583 TCAGCTTGCCTGTAAGTCCCGCAGTCAAGTGTGTACCTTCACAGGAGTGCCTGTGCG 1642
 Db 653 TCAG-CTGCGCTGTAAGTCCCGCAGTCAAGTGTGTACCTTCACAGGAGTGCCTGTGCG-- 707
 QY 1643 CAACCGAGCCAGGCGCTTCAGAGGAGCAG 1670
 Db 708 -CAACCGAGCCAGGCGCTTCAGAGGAGCAG 734

RESULT 14

AL602806 589 bp mRNA linear EST 14-AUG-2001
 LOCUS DKFZP686L2216_r1 686 (synonym: hlc3) Homo sapiens cDNA clone

DEFINITION DKFZP686L2216 5', mRNA sequence.
 ACCESSION AL602806
 VERSION AL602806.1 GI:15166312

KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 589)
 Mammalia: Eutheria: Primates: Catarrhini: Homiidae: Homo.
 Wambutt, R., Heubner, D., Mewes, W., Well, B. and Wiemann, S.
 EST (Wambutt, R., Heubner, D., Mewes, H.W., Well, B. and Wiemann, S.)
 Unpublished (1999)
 COMMENT Contact: Wambutt R

MIPS
 Am Kioferpitz 18a D-82152 Martinsried, Germany
 This is the 5' sequence of the clone insert

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
 sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
 consortium of the German Genome Project.

No sl sequence available.
 This clone (DKFZP686L2216) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

```

FEATURES
  source
    Location/Qualifiers
      1. 589
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /clone_id="686 (synonym: h1cc3)"
      /tissue_type="human skeletal muscle"
      /dev_stage="adult"
      /lab_host="DH10B"
      /note="Vector: pTRIPlex2; Site_1: SfiIA; Site_2: SfiIB;
      cDNA-collection"
BASE COUNT      113 a      181 c      203 g      91 t      1 others
ORIGIN
Query Match      23.1%; Score 586.4; DB 9; Length 589;
Best Local Similarity 99.7%; Pred. No. 3.7e-81;
Matches 587; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 71 GCTGCGGCGCTGCTCCATGGCCCTGTGTCACAGTGAAGCCGTTCCGCCCGGCGAGCGCG 130
    |||||||
Db 1 GCTGCGGCGCTGCTCCATGGCCCTGTGTCACAGTGAAGCCGTTCCGCCCGGCGAGCGCG 60

QY 131 CCTCCAGCGCCGCTGGGCGCTGGGACCAAGCGGTCACGCGAAGAGTCACTCCAGCGAA 190
    |||||||
Db 61 CCTCCAGCGCCGCTGGGCGCTGGGACCAAGCGGTCACGCGAAGAGTCACTCCAGCGAA 120

QY 191 GGCAGAGCTTGGCGGTGCTCCGTGGGGGCTGTCGAGGACGTCGAGGATGAGAGGGGCAATG 250
    |||||||
Db 121 GGCAGAGCTTGGCGGTGCTCCGTGGGGGCTGTCGAGGACGTCGAGGATGAGAGGGGCAATG 180

QY 251 ATGATGCAAGCAGAGCGCCAGTTCTGAGCCACAGAAAGAGCCCGCAGTAGAGAGAGCTCC 310
    |||||||
Db 181 ATGATGCAAGCAGAGCGCCAGTTCTGAGCCACAGAAAGAGCCCGCAGTAGAGAGAGCTCC 240

QY 311 ACGGGGACCAACAGACTTCGCGGCAAGGATCCAGAGTCCCGCAGAGCAGAGAGCAGA 370
    |||||||
Db 241 ACGGGGACCAACAGACTTCGCGGCAAGGATCCAGAGTCCCGCAGAGCAGAGAGCAGA 300

QY 371 GGCAGACCTGCTACCTCATGTGTACAGTGTGAGGCGCCAGATGACATCCGCGCGGCGAG 430
    |||||||
Db 301 GGCAGACCTGCTACCTCATGTGTACAGTGTGAGGCGCCAGATGACATCCGCGCGGCGAG 360

QY 431 CCCAGCTGAGAGCACCCTGCGGCTCCCGGCTCCGCTACCTGCTGTAATTTCTACACAG 490
    |||||||
Db 361 CCCAGCTGAGAGCACCCTGCGGCTCCCGGCTCCGCTACCTGCTGTAATTTCTACACAG 420

QY 491 AAGAGAAAGTCTGAGCCAGATGAGACGTCCTCTGCGGTGATTTCCCTGACAGCA 550
    |||||||
Db 421 AAGAGAAAGTCTGAGCCAGATGAGACGTCCTCTGCGGTGATTTCCCTGACAGCA 480

QY 551 GCTGCCCCAGCTGACCCCTGGGCTGCTGCCCCCTGTGAGTGAACCCAGGCTACT 610
    |||||||
Db 481 GCTGCCCCAGCTGACCCCTGGGCTGCTGCCCCCTGTGAGTGAACCCAGGCTACT 540

QY 611 TAGATGAGACGGGGGCTTACGCTGACGTCGTGGGCAAAAGCCGAGT 659
    |||||||
Db 541 TAGATGAGACGGGGGCTTACGCTGACGTCGTGGGCAAAAGCCGAGT 589

RESULT 15
LOCUS      BG831015      672 bp      mRNA      linear      EST 22-MAY-2001
DEFINITION 602766736F1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:4908729 5',
            mRNA sequence.
ACCESSION  BG831015
VERSION    BG831015.1 GI:14178602
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 672)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
```

```

TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished (1999)
COMMENT    Contact: Robert Strausberg, Ph.D.
            Email: gcaps-remail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNLN at:
            http://image.lnl.gov
            Plate: LCM1815 row: h column: 10
            High quality sequence stop: 670.

FEATURES
  source
    Location/Qualifiers
      1. 672
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /clone_id="IMAGE:4908729"
      /clone_id="NIH_MGC_42"
      /tissue_type="epithelioid carcinoma cell line"
      /lab_host="DH10B (phage-resistant)"
      /note="Organ: pancreas; Vector: pOT7; Site_1: XhoI;
      Site_2: EcoRI; cDNA made by oligo-dT priming.
      Directionally cloned into EcoRI/XhoI sites using the
      following 5' adaptor: GGCAGAG(G). Size-selected >500bp
      for average insert size 1.8kb. Library constructed by Ling
      Hong in the laboratory of Gerald M. Rubin (University of
      California, Berkeley) using ZAP-cDNA synthesis kit
      (Stratagene) and Superscript II RT (Life Technologies).
      Note: This is a NIH_MGC Library. 1"
BASE COUNT      146 a      219 c      200 g      107 t
ORIGIN
Query Match      22.4%; Score 569; DB 12; Length 672;
Best Local Similarity 98.5%; Pred. No. 1.7e-78;
Matches 606; Conservative 0; Mismatches 5; Indels 4; Gaps 3;

QY 1301 CGGCCAGAACTGAGGGGTGGTGGGAGAGAAAGTTGAGGCATGGAAGAGCCAGCA 1360
    |||||||
Db 61 CGGCCAGAACTGAGGGGTGGTGGGAGAGAAAGTTGAGGCATGGAAGAGCCAGCA 120

QY 1361 GCGCCGAAAGAAAGCCCTGGGGGCGACGATATAACCTCCGAGGGTCATGAGT 1420
    |||||||
Db 121 GCGCCGAAAGAAAGCCCTGGGGGCGACGATATAACCTCCGAGGGTCATGAGT 178

QY 1421 CCATCAGCTTCTGAGAGCCCTCTCTTGGGAGCTGGAGACCACTCATGACATGACAT 1480
    |||||||
Db 179 CCATCAGCTTCTGAGAGCCCTCTCTTGGGAGCTGGAGACCACTCATGAGAT 236

QY 1481 GCGAGAGCTCTTCTCTTCCCAAGAGTCTTCAATGAAGACCTGCGAGCCCTCCACA 1540
    |||||||
Db 237 GCGAGAGCTCTTCTCTTCCCAAGAGTCTTCAATGAAGACCTGCGAGCCCTCCACA 296

QY 1541 GCTTGCAAGACCAAGAGGAGCGACAGTGAAGAGGGGCGCTCAGCTCCCTGAAGTC 1600
    |||||||
Db 297 GCTTGCAAGACCAAGAGGAGCGACAGTGAAGAGGGGCGCTCAGCTCCCTGAAGTC 356

QY 1601 CCGCCAGTCAAGTGTACCTTCAGAGGAGTGCCTGTGGCCAAACCGGACAGGCTT 1660
    |||||||
Db 357 CCGCCAGTCAAGTGTGTACCTTCAGAGGAGTGCCTGTGGCCAAACCGGACAGGCTT 416

QY 1661 CCAGAGCAGAGCAGAGGGGCGAGGGGCGAGGAGAGAGCCCTGATTTCTCTAGCC 1720
    |||||||
Db 417 CCAGAGCAGAGCAGAGGGGCGAGGGGCGAGGAGAGAGCCCTGATTTCTCTAGCC 476

QY 1721 CAGGTTCCGAAGGTGTGAGACAGGCGCAGTGCATATGAGAGAGAGGCGAGGC 1780
    |||||||
Db 477 CAGGTTCCGAAGGTGTGAGACAGGCGCAGTGCATATGAGAGAGAGGCGAGGC 536

QY 1781 CTGAGCCCTCACACATGGCCAGCTCCCTTACACACTGAAGAGATCCACACATCTTGA 1840
    |||||||
Db 537 CTGAGCCCTCACACATGGCCAGCTCCCTTACACACTGAAGAGATCCACACATCTTGA 596
```

Mon Apr 14 12:02:12 2003

us-09-761-640-3.rst

Page 12

[illegible]

Search completed: April 11, 2003, 23:49:11
Job time : 6920.21 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 11, 2003, 03:06:43 ; Search time 7412 Seconds

(without alignments)
1849.355 Million cell updates/sec

Title: US-09-761-640-4

Perfect score: 2436
Sequence: 1 MALVTYSRSPGSGASTPVG.....PNEGFLROLQYQGLTART 471

Scoring table:

BLOSUM62	
Xgapop 10.0 , Ygapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlh
-O=/cgnr_1/USPRO.spool/us09761640/runat_08042003.141434.20357/app-query.fasta.1.647
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=Dlosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=Pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=us09761640.&CGN.1.1.4921.&runat.08042003.141434.20357 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEROJURY -NEG_SCORES=0 -NAIT -LONGLOG -DEV TIMEOUT=120
-MARR_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl: *
1: gb_ba: *
2: gb_hg: *
3: gb_in: *
4: gb_on: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pt: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vi: *
15: em_da: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_com: *
21: em_or: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sts: *
28: em_un: *

29: em_vi: *
30: em_htg_hum: *
31: em_htg_inm: *
32: em_htg_other: *
33: em_htg_mus: *
34: em_htg_pln: *
35: em_htg_rtd: *
36: em_htg_mam: *
37: em_htg_vrt: *
38: em_sy: *
39: em_htgo_hum: *
40: em_htgo_mus: *
41: em_htgo_other: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query % Match	Length	DB ID	Description
1	2436	100.0	1416	6	AX451362 Sequence
2	2426	99.6	1980	6	AX451343 Sequence
3	2426	99.6	2781	6	AX099939 Sequence
4	2426	99.6	2808	9	AK074432 Sequence
5	2392	98.2	1416	9	AB072360 Homo sapi
6	2333	95.8	2905	9	AK094226 Homo sapi
7	1991.5	81.8	2736	10	BC028922 Mus muscu
8	1987.5	81.6	2322	6	AX398774 Sequence
9	1470.5	60.4	2064	6	AX405828 Sequence
10	1470.5	60.4	2064	9	AK000522 Sequence
11	1405.5	57.7	160963	2	AP002776 Homo sapi
12	1405.5	57.7	171566	2	AP001885 Homo sapi
13	1113	45.7	2280	9	AB072355 Homo sapi
14	1113	45.7	3817	9	AB072355 Homo sapi
15	1113	45.7	6374	6	AX180874 Sequence
16	1094.5	44.9	1796	9	AB072358 Homo sapi
17	1077	44.2	2433	9	AK095421 Homo sapi
18	1076.5	44.2	2260	6	AX180876 Sequence
19	1045	42.9	1711	6	AX223964 Sequence
20	1004.5	41.2	1771	6	AX202239 Sequence
21	994	40.8	1905	9	BC004176 Homo sapi
22	994	40.8	1905	9	BC004210 Homo sapi
23	980	40.2	1949	6	AX223960 Sequence
24	918.5	37.7	4417	3	AB036834 Drosophi1
25	852	35.0	3488	6	AX406972 Sequence
26	793	32.6	1026	6	AX086034 Sequence
27	760.5	31.2	1755	6	AX398768 Sequence
28	760.5	31.2	1755	9	AK001790 Homo sapi
29	733	30.1	181132	3	AC008206 Drosophi1
30	733	30.1	227398	3	AE003750 Drosophi1
31	699	28.7	113474	2	AC013932 Drosophi1
32	689	28.3	599	6	AX398772 Sequence
33	687	28.2	409	6	AX398771 Sequence
34	591.5	24.3	426	6	AX398765 Sequence
35	589	24.2	1052	6	AX223966 Sequence
36	426	17.5	5463	9	AB037719 Homo sapi
37	413	17.0	494	6	AX398773 Sequence
38	405	16.6	386	9	AF484838 Homo sapi
39	378	15.5	191571	2	AL845302 Dario rer
40	366.5	15.0	957	9	BC008941 Homo sapi
41	359	14.7	1008	9	AB072359 Homo sapi
42	322	13.2	935	9	BC011636 Homo sapi
43	310.5	12.7	1031	9	AB072357 Homo sapi
44	287	11.8	1083	8	AT118620 Arabidopsi
45	271	11.1	141017	2	AC116962 Dictyoste

ALIGNMENTS

RESULT 1

AX451362 1416 bp DNA Linear PAT 03-JUL-2002
LOCUS AX451362
DEFINITION Sequence 20 from Patent WO0224740.
ACCESSION AX451362
VERSION AX451362.1 GI:21698396
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 Luche, R.M. and Wei, B.
AUTHORS Dsp-15 dual-specificity phosphatase
TITLE Patent: WO 0224740-A 20 28-MAR-2002;
JOURNAL Ceptyr, Inc. (US)
FEATURES
source 1. 1416
/organism="Mus musculus"
/db_xref="taxon:10090"
BASE COUNT 288 a 450 c 449 g 229 t
ORIGIN
Alignment Scores:
Pred. No.: 3,88e-165 Length: 1416
Score: 2436.00 Matches: 471
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
US-09-761-640-4 (1-471) x AX451362 (1-1416)
QY 1 MetAlaLeuValThrValSerArgSerProGlySerGlyAlaSerThrProValGly 20
DB 1 ATGGCCCTGGTACAGTGAAGCCGTTGCCCGGCGAGCGGCGCTCCAGCCGGCGGGG 60
QY 21 ProTyrPAspGlnAlaValGlnArgArgSerArgLeuGlnArgGlnSerPheAlaVal 40
DB 61 CCTGGGAGACAGGCGGTCCAGGAGAGAGAGTGCAGCTCCAGCGAAGGCGAGCTTGGCGGTG 120
QY 41 LeuAlaGlyAlaValLeuGlyLeuGlnAspGlyGlyAspAsnAspAspAlaAlaGluAla 60
DB 121 CTCCTGGGGGCGTGTCTGGAGTGCAGATGAGAGGAGGAGCAATGATGATGCAGCAGAGGCC 180
QY 61 SerSerGlnProThrGlnLysAlaProSerGlnGluGluLeuHisGlyAspGlnThrAsp 80
DB 181 ATTTCTGAGCCACAGAGAGAGGCCCGGAGTGAAGAGAGACTCCAGCGGAGACAGACAGAC 240
QY 81 PheGlyGlnGlySerGlnSerProGlnLysGlnGluGlnArgGlnHisLeuHisLeu 100
DB 241 TTCGGGCAAGGATCCAGAGTCCCGAGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
QY 101 MetValGlnLeuLeuArgProGlnAspAspIleArgLeuAlaAlaGlnLeuGluAlaPro 120
DB 301 ATGGATACACTCTGTGAGGCGGAGAGATGACATCCGCTGGCAGCCCGAGCTGGAGGAGCC 360
QY 121 ArgProProArgLeuArgTyrLeuLeuValValSerThrArgGluGlyGlyLeuSer 140
DB 361 CGGCTCCCGGCGGCTCCGCTACTGCTGATGTTCTACACAGAGAGAGAGAGAGAGAGAGAG 420
QY 141 GlnAspGluThrValLeuLeuGlyValAspPheProAspSerSerSerProSerCysThr 160
DB 421 CAGGATGAGACAGGCTCTCTGGGCGGTGATTTCCCTGACAGAGAGCTCCCGGAGAGCTCAC 480
QY 161 LeuGlyLeuValLeuProLeuTyrPAspThrGlnValTyrLeuAspGlyAspGlyGly 180
DB 481 CTGGGCTGCTGTGGCTCTGGAGTGAACCCAGAGAGTACTTAATGAGAGAGCGGGGCG 540
QY 181 PheSerValThrSerGlyGlnSerArgIlePhePheLysProIleSerIleGlnThrMet 200
DB 541 TTCAGGCTGACGTGTGTGGGCAAGCCGAGATCTTAAGAGAGAGAGAGAGAGAGAGAGAG 600
QY 201 TyrPalaThrLeuGlnValLeuHisGlnAlaCysGluAlaAlaLeuGlySerGlyLeuVal 220

DB 601 TGGGCGACACTCCAGATATGCACAGAGATGAGGAGAGCTCTAGGAGCGGCTGTGA 660
QY 221 ProGlyGlySerAlaLeuThrTyrPalaSerHisTyrGlnGluArgLeuAsnSerGluGln 240
DB 661 TTTTGGTGGAGAGGCCCTCACTCCGCGGAGCCCTCAGAGAGAGAGAGAGAGAGAGAGAGAG 720
QY 241 SerCysLeuAsnGluThrThrAlaMetAlaAspLeuGlnGluSerLeuArgProProSerAla 260
DB 721 AGCTGCTCATGATGATGAGAGCGGCTATGGCCGAGCTCGAGATCTCTCGGCTCCAGCCGC 780
QY 261 GluProGlyGlySerSerGlnGlnGlnGlnMetGlnGlnAlaIleArgAlaGluLeuThr 280
DB 781 GAGCTGGCGGCTCTCTCAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 281 LysValLeuAspValSerAspLeuGluSerValThrSerLysGluThrArgGlnAlaLeu 300
DB 841 AAGGTGTGGATGTCACTGACCTGAGAGAGTGTCACTTCCAAAGAGATCGCGGAGGCTCTG 900
QY 301 GluLeuArgLeuGlyLeuProLeuGlnGlnTyrArgAspPheIleAspAsnGlnMetLeu 320
DB 901 GAGCTGGCGCTGGGCTCCCTCCAGAGAGTACCTGAGCTTATGACAGACAGATGCTG 960
QY 321 LeuLeuValAlaGlnArgAspArgAlaSerArgIlePheProHisLeuThrLeuGlySer 340
DB 961 CTGCTGTGGACACAGCGGAGCCAGAGCTCCGCAATCTTCCCGACCTGACTCTGAGCTGCA 1020
QY 341 GluTyrPAsnAlaAlaAsnLeuGlnGluLeuGlnArgAsnArgValThrHisTyrLeuAsn 360
DB 1021 GAGTGGAGACAGCAAACTGGAGAGAGCTGCAGAGAGAACAGGGGTCAACCACTTGTGAC 1080
QY 361 MetAlaArgGluIleAspAsnPheTyrProGluArgPheThrTyrHisAsnValArgLeu 380
DB 1081 ATGGCCCGGAGATTTGACACTTACCTTGAGCGCTTACCTACACAGATGTGGCGCTC 1140
QY 381 TyrPAspGluSerAlaGlnLeuLeuProHisTyrPlyGluThrHisArgPheIleGlu 400
DB 1141 TGGGATGAGAGTCCGCGCAGCTGCTGCGCACTGGAGAGAGACACAGCGCTTCAATTGAG 1200
QY 401 AlaAlaArgAlaGlnGlyThrHisValLeuValHisCysLysMetGlyValSerArgSer 420
DB 1201 GCTGCAGAGACCAAGGAGCCACCACTGCTGTGTCACATGCAATGAGGCGCTCAGCGGCTCA 1260
QY 421 AlaAlaThrValLeuAlaValAlaMetLysGlnTyrGluCysSerLeuGlnAlaLeu 440
DB 1261 GCGGCGACAGCTGTGGCTATGCCATGAAGCAGTACGATGACAGCTGGAGAGCGGCTG 1320
QY 441 ArgHisValGlnGluLeuArgProIleAlaArgProAsnProGlyPheLeuArgGlnLeu 460
DB 1321 CGCCAGGTGCAGAGAGCTCCGCGCCATCCGCGCCCAACCCCTGGGCTTCTGTGCGCAGCTG 1380
QY 461 GlnIleTyrGlnGlyTyrLeuThrAlaArgThr 471
DB 1381 CAGATCTACAGAGGCACTCTGACGCGCCAGAACCC 1413
RESULT 2
AX451343 1980 bp DNA Linear PAT 03-JUL-2002
LOCUS AX451343
DEFINITION Sequence 1 from Patent WO0224740.
ACCESSION AX451343
VERSION AX451343.1 GI:21698394
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 Luche, R.M. and Wei, B.
AUTHORS Dsp-15 dual-specificity phosphatase
TITLE Patent: WO 0224740-A 1 28-MAR-2002;
JOURNAL Ceptyr, Inc. (US)
FEATURES
source 1. 1980
Location/Qualifiers

QY 1 MetAlaLeuValThrValSerArgSerProGlySerGlyAlaSerThrProValGly 20
 DB 87 ATGGCCCTGGTACATGAGCGCTTGGCCCCGGGAGCGGCGCTCCAGCGCCGTTGGGG 146
 QY 21 ProThrAspGlnAlaValGlnArgArgSerArgLeuGlnArgArgGlnSerPheAlaVal 40
 DB 147 CCTTGGGACAGCGGTGTCAGCGCAAGAGTCTCCAGGAGGAGGAGAGAGCTTTGGGCTG 206
 QY 41 LeuArgGlyAlaValLeuGlyLeuGlnAspGlyGlyAspAsnAspAlaAlaGlnAla 60
 DB 207 CTCGCGGGGGGCTCTCTGGGACGTCAGAGATGAGGGGAGACATATATATGTCAGAGAGCC 266
 QY 61 SerSerGluProThrGluValAspAlaProSerGluGluGluGluGlnArgGlnHisLeu 80
 DB 267 AGTTTGAGACCAACAGAGAAGGCCCGGAGTGAAGAGAGAGCTCCACGGGGAGCAGACAGAC 326
 QY 81 PheGlyGlnGlySerGlnSerProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 100
 DB 327 TTGCGGCAAGAGATCCAGAGTCCCGAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 386
 QY 101 MetValGlnLeuLeuValArgProGlnAspAlaArgLeuAlaAlaGlnLeuGlnAlaPro 120
 DB 387 ATGTCACAGCTGCTGAGCGCGGAGATGACATCCGCTGGAGCCCAAGCTGGAGGACACC 446
 QY 121 ArgProProArgLeuArgGlyLeuLeuValValSerThrArgGlnGlnGlnGlnGlnSer 140
 DB 447 CGGCTCCCGGCGCTCGTACCTGCTGATGTTTCTACAGAGAGAGAGAGAGAGAGAGAGAG 506
 QY 141 GlnAspGlnThrValLeuGlnGlyValAspPheProAspSerSerSerSerSerSerSer 160
 DB 507 CAGAGTGAAGAGCGTCTCTCGGGCGTGAGTTCCTCCAGACGCGTCCCGACACTCCAC 566
 QY 161 LeuGlyLeuValLeuProLeuThrSerAspThrGlnValThrLeuAspGlyAspGlyGly 180
 DB 567 CTGGGCTGCTGTTGGCCCTCTGGAGTGCACACCCAGTGTACTATGAGAGCGGGGGC 626
 QY 181 PheSerValThrSerGlyGlnSerArgGlnPheLeuSerProLeuSerLeuGlnThrMet 200
 DB 627 TTCAGCGTCAAGTCTGCTGGGCAAGCCGAGTCTTCAAGCCCTCCATCCAGACCATG 686
 QY 201 ThrAlaThrLeuGlnValLeuHisGlnAlaCysGlnAlaAlaLeuGlnSerGlyLeuVal 220
 DB 687 TGGGCCACACTCCAGATATTGCACCAAGATGAGGAGCTTAGCGAGCGGCTTGTGA 746
 QY 221 ProGlyGlySerAlaLeuThrThrAlaSerHisTyrGlnGlnArgLeuAsnSerGlnGln 240
 DB 747 CCGGGTGGAGTCCCTCACTACCTGGCCAGCACTACCAAGAGAGACTGAAGTCCGAACAG 806
 QY 241 SerCysLeuAsnGlnThrPheAlaMetAlaAspLeuGlnSerLeuArgProProSerAla 260
 DB 807 AGCTGCTCAATGATGAGACGCGCTATGGCCGACTGAGTCTCTGGCGCTCCCACTGGCC 866
 QY 261 GluProGlyGlySerSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 280
 DB 867 GAGCTGGGGGCTCTCTCAACAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 926
 QY 281 LysValLeuAspValSerAspLeuGlnSerValThrSerLysGlnGlnGlnGlnGlnGln 300
 DB 927 AAAGTGTGGAGTGCAGTACCTGGAGAGTGTCACTTCCAAAGAGAGTCCGCCAGGCTCTG 986
 QY 301 GluLeuArgLeuGlyLeuProLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 320
 DB 987 GAGCTGGCGCTGGGCTCCCTCCAGCAGTACCGTCACTTCATCGACAACAGATGCTG 1046
 QY 321 LeuLeuValAlaGlnArgArgArgAlaSerArgGlnPheProHisLeuThrLeuGlySer 340
 DB 1047 CTGCTGGTGGCAACAGGGAGCCGAGCTCCCGATCTTCCCACTTACCTGGGCTCA 1106
 QY 341 GluThrAspAlaAlaAsnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 360
 DB 1107 GAGTGAAGCGAACAACCTGCGAGAGCTGCAGAGAGACAAGGTCCACCATCTTGAAC 1166
 QY 361 MetAlaArgGlnLeuAspAsnPheThrProGluArgPheThrHisAsnValArgLeu 380

DB 1167 ATGGCCGGGAGATGACAACCTTACCTGAGCGGCTTCACTACCAACAATGAGCGCTTC 1226
 QY 381 ThrAspGluGluSerAlaGlnLeuLeuProHisTyrGlnGlnGlnGlnGlnGlnGlnGln 400
 DB 1227 TGGGATGAGAGTGGCGCCAGCTGCTGCCGACTGGAAGAGAGAGAGAGAGAGAGAGAGAG 1286
 QY 401 AlaAlaArgAlaGlnGlyThrHisValLeuValHisCysLysMetGlyAlaSerArgSer 420
 DB 1287 GCTGCAAGAGACAGGAGCCACCCAGCTGCTGCTCACTGCAAGATGGGCGCTGAGCCCTTA 1346
 QY 421 AlaAlaThrValLeuAlaTyrAlaMetLysGlnTyrGluCysSerLeuGlnAlaLeu 440
 DB 1347 GCGCCACAGATGCTGCTGCTATGCTATGCAAGAGATGAGTACCTGGAGCAGGCGCTG 1406
 QY 441 ArgHisValGlnGluLeuArgProLeuAlaArgProAspProGlyPheLeuArgGlnLeu 460
 DB 1407 CGCCAGTGCAGAGAGCTCGGCGCCATCGCCGCCACACCTGCTTCTGGCGCAGCTG 1466
 QY 461 GlnIleTyrGlnGlyIleLeuThrAla 469
 DB 1467 CAGATCTACAGGAGCATCTCTGACGCGC 1493
 RESULT 4
 AK074432
 LOCUS
 DEFINITION
 AK074432
 AK074432.1 GI:18677033
 VERSION
 AK074432.1 GI:18677033
 KEYWORDS
 Homo sapiens signet-ring cell carcinoma cell_line:KATO III cDNA to mRNA, clone_lib:KAT clone:KAT12021.
 SOURCE
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS
 1 Watanabe,K., Kumagai,A., Itakura,S., Yamazaki,M., Tashiro,H., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
 NEDO human cDNA sequencing project
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 2 (bases 1 to 2808)
 Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T., Shibahara,T., Tanaka,T. and Nakamura,Y.
 Direct Submission
 Submitted (14-FEB-2002) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
 (E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)
 COMMENT
 NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan: cDNA full insert sequencing: Research Association for Biotechnology: cDNA library construction: 5'-& 3'-end one pass sequencing: Department of Vitrology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).
 location/Qualifiers
 1, 2808
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="KAT12021"
 /cell_line="KATO III"
 /cell_type="signet-ring cell carcinoma"
 /clone_lib="KAT"
 /note="cloning vector: pME18SFL3"
 97, 2076
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="BAB85080.1"
 /db_xref="GI:18677034"
 /translation="MALYTVRSPPGCSASIPVGPMDQAVQRSRRLQRQSFAYLRGA VLGIDGDNDNDAAEASSEPEKAPSEELHGDGTDREQSGSQSPQKEDEQRQHLHLVY QLRPQDILRLAQLLEAPRPRLRLVTVSTREEGLSQDETVLLGLVDFPSSPSCV

LGVLTPLMSDPQVYVLDGDDGGSVTSGGGSRIFKPSITQMTATLQVHQCCEALIT
 LVEGSLALTWASHYQERLUNDSBOCLNETATADLEISLTMPBAEPGGSECOMQALIT
 ALMKVLDVSDLESVTSKEITROALELRIGTLPLOOYREFIDNOMILLVAQRASRIFF
 HLTVSEEMNANLEELORNRVTHILNARLEIDNYPPEFTYHNRILMDEESAOLPLPH
 KETHRTLEAARAGOSTHYLVHCKMGVSSAAVLYAMQVCSYALALRHQOELRPTV
 RPNGFTLRDOLDITOGILTLASRQSHVMECKGCVSPDEHPDAEVSTPEPFLPEDEGGG
 EEKVLVMEESQAPAKEEPGRPRINLRQPMKSSILDELTLESTSTMDPEVFSISL
 EESSHEEPLQEPOLARTKGGQOVDRCGRPLVMSSTVLTGSAVVAHQTAQFOEJEE
 GGQGGGQGPCSTSTPRSRKVVYRQASVHDSGSEGEA"

Alignment Scores:	
Pred. No.:	4,66e-164
Score:	2426.00
Percent Similarity:	100.00%
Best local Similarity:	100.00%
Query Match:	99.59%
DB:	9
Gaps:	0
Length:	2808
Matches:	469
Conservative:	0
Mismatches:	0
Indels:	0

US-09-761-640-4 (1-471) x AK074432 (1-2808)

Oy	1	MelLaleuVal1ThrValSerArgSerProProGlySerGlyAlaSerThrProValGly	20
Db	97	ATGGCCCTGGGTACAGTAgAGCCGTTCCCCCGGGCAGGGCCCTTCACAGCCCGTGGG	156
Oy	21	ProTrpAspGlnAlaValAlaGlnArgArgSerArgLeuGlnInArgArgGlnSerPheAlaVal	40
Db	157	CCCTGGAGCACAGGGCGTCTCCAGCAAGAGTGCAGTCCAGGCAAGCGACCTTTGGCGT	216
Oy	41	LeuArgGlyAlaValLeuGlyLeuGlnAspGlyGlyAspAsnAspAspAlaAlaGlyAla	60
Db	217	CTCCGTGGGGCTGTCTCTGGAGCTCAGAGATGGAGGGAGCATGTATGTCAGCAGAGGCC	276
Oy	61	SerSerGluProThrGluValAlaProSerGlnGlnGluLeuHisGlyAspGlnThrAsp	80
Db	277	AGTTTGTGGCCACACAGAGAAAGCCCCAGATGAAGAGAGCTCCACGGGGACCAAGACAGAC	336
Oy	81	PheGlyGlnGlySerGlnSerProGlnIleValGlnGlnGluGlnInArgGlnHisLeuHisLeu	100
Db	337	TTTCGGGCAAGGATCTCCAGAGTCCCCAGAGAGGAGGAGGAGGAGCGACACTGCAGCTC	396
Oy	101	MetValGlnLeuLeuArgArgProGlnAspAspIleArgLeuAlaAlaGlnLeuGlnAlaPro	120
Db	397	ATGGACAGCTGTGAGGCCGACAGAGATGACATCCGCTGGACGCCAGCTGAGGACACC	456
Oy	121	ArgProProAlaGlyLeuArgGlyLeuLeuValValSerThrArgGlnGlyGlnGlyLeuSer	140
Db	457	CGGCTTCCCGGGCTCCGCTACCTCTGGTATTTCTACAGAGAGAGAGAGAGGCTGTAGC	516
Oy	141	GlnAspGlnThrValLeuLeuGlnGlyValAspPheProAspSerSerSerProSerCysThr	160
Db	517	CAGGATGAGAGCGTCTCTCTGGCGGTGATTTCCCTCAGACAGACTCTCCACACTGCACC	576
Oy	161	LeuGlyLeuValLeuProLeuTrpSerAspThrGlnValArgLeuAspGlyAspGlyGly	180
Db	577	CTGGGCTGTGCTTGGCCCTCTGGAGTGCACCCAGGCTGACTAGTAGAGAGAGGGGGC	636
Oy	181	PheSerValThrSerGlyGlyGlnSerArgIlePheIleProIleSerIleGlnThrMet	200
Db	637	TTTCACCGTAGCGTGTGGTGGGCAAGCGGATCTTTCAGGCCCATCTTCATCCAGACCATG	696
Oy	201	TrpAlaThrLeuGlnValLeuHisGlnAlaCysGlnAlaAlaLeuGlnGlySerGlyLeuVal	220
Db	697	TGGGCCACACTCCAGGATTTGCACCAAGCATGTAGGACAGCTTAGGGCAGCGGCTTGTA	756
Oy	221	ProGlyGlySerAlaLeuThrTrpAlaSerHisTyrGlnGlnIleArgLeuAsnSerGlyGln	240
Db	757	CCGGGTGGCAGAGCCCTCACCCTGGGGCCAGCCATACCAAGAGAGAGACTCGAATCCGAACAG	816
Oy	241	SerCysLeuAsnGluTrpThrAlaMetAlaAspLeuGlnSerLeuAlaArgProProSerAla	260
Db	817	AGCTTCCTCAATGATGGAGCGCTATAGGCGACCTGTGAGTCTTGCGCGCTCCACAGGCC	876

OY	261	GIUProGlyGlySerSerGluGlnGlnGlnMetGluGlnAlaIleArgAlaGluLeu	280
Db	877	GAGCCTGGCGGGCTCTCCAGAACAGGAGAGATGGAGCGGCATCCGGCTGAGCTGG	936
OY	281	LysValIleAspValSerAspLeuGluSerValThrSerIysGluIleArgGlnAlaLeu	300
Db	937	AAAGCTTGAGATGTCAGTGACCTGGAGAGTGCACCTCCAAAGAGATCCGCAGGCTCTG	996
OY	301	GluLeuArgLeuGlyLeuArgProLeuGlnGlnIleThrArgAspPheIleAspGlnMetLeu	320
Db	997	GAGCTGCGCCCTGGGGCTCCCTCCAGAGATACCTGATCTTCATCGAACACAGATGCTG	1056
OY	321	LeuLeuValAlaGlnArgAspArgIleSerArgIlePheProHisLeuThrLeuGlySer	340
Db	1057	CTGCTGGTGAGCACAGCGGAGCCGAGACCTCCCGCATCTTCCCCCACTCTTACTGGGCTCA	1116
OY	341	GluTrpAsnAlaIleAsnLeuGlnGlnLeuGlnGlnArgAsnArgValThrHisIleLeuAsn	360
Db	1117	GAGTGGAAACGCGACAAACCTGGAGAGCTGCAGAGGAAACAGGCTACCCACATCTTGAAAC	1176
OY	361	MetAlaIleArgGluIleAspAsnPheArgProGluArgPheThrThrHisAsnValArgLeu	380
Db	1177	ATGGCCCGGGAGATTGACAACTTACCTCGAGCCCTTCACCTACCAACAATGGCGCTC	1236
OY	381	TrpAspGluGluSerAlaGlnLeuLeuProHisTrpIysGluThrHisArgPheIleGlu	400
Db	1237	TGGGATGAGAGATCGGGCCAGCTGCTGCCCGCACTGGAGAGAACGCCACCGCTTCAATTAG	1296
OY	401	AlaAlaIleArgAlaGlnGlyThrHisValLeuValHisCysLysMetGlyValSerArgSer	420
Db	1297	GCTGCAGAGGACACAGGAGCACCCAGCTGGTGCCACTGCAGAAATGGGGTCCAGCGCTCA	1356
OY	421	AlaAlaIleThrValLeuAlaIleThrAlaMetIysGlnIleArgIleCysSerLeuGlnGlnAlaLeu	440
Db	1357	GGGGCCACAGTCTGGCTTATGCCATGAGACAGTACGAAATGCAAGCTGGAGCAGGCCCTG	1416
OY	441	ArgHisValGlnGlnLeuArgProIleAlaArgProAsnProGlyPheLeuArgGlnLeu	460
Db	1417	CGCCACGTGCAGAGATCGCGCCATCGGCCCGCCCAACCCCTGGCTCTCGTGGCCAGCTG	1476
OY	461	GlnIleThrGlnGlnIleLeuThrAla 469	
Db	1477	CAGATCTACAGGCACTCTGACGGCC 1503	

RESULT

LOCUS	AB072360	1416 bp	mRNA	linear	PRI 29-JAN-2002
DEFINITION	Homo sapiens mRNA for hssh-3, complete cds.				

ACCESSION AB072360
 VERSION AB072360.1 GI:18376668
 KEYWORDS
 SOURCE Homo sapiens cDNA to mRNA

REFERENCE
1
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi,
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

AUTHORS Niva, R., Nagata-Onashi, K., Taketuchi, M., Mizuno, K., and Uemura, T
TITLE Control of actin reorganization by Slingshot, a family of phosphatases that dephosphorylate ADF/cofilin
JOURNAL Cell 108 (2), 233-246 (2002)

MEDLINE	21822082
REFERENCE	2 (bases 1 to 1416)
AUTHORS	Niwa, R., Nagata-Ohashi, K., Hay, B. A., Takeichi, M., Mizuno, K. and Uemura, T.

TITLE Direct Submission
JOURNAL Submitted (29-SEP-2001) Tadashi Uemura, Laboratory of Molecular Genetics, The Institute for Virus Research, Kyoto University; Shogoin-Kawara-cho 53, Sakyo-ku, Kyoto 606-8507, Japan

FEATURES	Location/Qualifiers
source	1. .1416

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 1. .1416
 /gene="hSSH-3"
 1. .1416
 /gene="hSSH-3"
 /standard_name="slingshot-3"
 /note="a putative dual specific phosphatase"
 /codon_start=1
 /product="hSSH-3"
 /protein_id="BAB84119.1"
 /db_xref="GI:18376669"
 /translation="MALVYSGSPGSGASTPVGFMDOAVQRRSLRQSRQSFAYLRGA
 VLGDGDNDNDVAVSESPTEKASPEELHGDQDFQSGSQSPQKQEOHQLHLMY
 QLRPDDIRLAQLEAPRPRIYLVSTRREGSLGSDFTVLGDFPDSSPSC
 LGVLPLWSDTVYLDGDGFGSVTSGGGSRIFKPSIQTMATLOVLHQACEALGSG
 LVPGSALTWASHYQERLNSQSLNEMTMADLESRLPRSPKSGSEODMEQAIR
 AEIMKLVSDLESVTSKEIRQLMSGAMSPSSSTVTSSTRCCMHMSQDRASIFP
 HLYLSEWMAALELQRRNRYTHILNMAREDINFPERTYHNHVLMEESQQLPFA
 KTHRFIEARAGQTHVHVCNMGVSRSAATVLAAMQYECSEQALRHVQELRPIA
 RNPSPRLQIYQGLIAR"
 RNPSPRLQIYQGLIAR"

BASE COUNT 288 a 449 c 449 g 230 t
 ORIGIN

Alignment Scores:

Pred. No.: 5, 39e-162 Length: 1416
 Score: 2392.00 Matches: 469
 Percent Similarity: 99.36% Conservative: 0
 Best Local Similarity: 99.36% Mismatches: 2
 Query Match: 98.19% Indels: 0
 Gaps: 0

US-09-761-640-4 (1-471) x AB072360 (1-1416)

Oy 1 MetAlaLeuValThrValSerArgSerProGlySerGlyAlaSerThrProValGly 20
 Db 1 ATGCCCTGGTGCACAGTGAAGCCGTTCCGCCCGGCGACGGCTCCACGCCCTGGGG 60
 Oy 21 ProTrpAspGlnAlaValGlnArgSerArgLeuGlnArgGlnSerPheAlaVal 40
 Db 61 CCGTGGGACGAGGGGCTCCAGAGAGTGCCTCCAGCCAGGCAAGCCTTGGGCTG 120
 Oy 41 LeuArgGlyAlaValLeuGlyLeuGlnAspGlyAlaAspAlaAlaGlnAla 60
 Db 121 CTCCCTGGGGCTGCTCTGCTGAGTGAAGGAGGCAATGATGATGAGTGAAGGCC 180
 Oy 61 SerSerGluProThrGlnLysAlaProSerGluGluGluLeuHisGlyAspGlnThr 80
 Db 181 AGTTCTGAGCCAAAGAGAGAGCCCGAGTGAAGGAGCTCCAGGGAGCCAGACAG 240
 Oy 81 PheGlyGlnGlySerGlnSerProGlnLysGlnGluGlnArgGlnHisLeuHisLeu 100
 Db 241 TTCGGGCAAGGATCCAGAGTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300
 Oy 101 MetValGlnLeuLeuArgProGlnAspAspLeuArgLeuAlaAlaGlnLeuAlaPro 120
 Db 301 ATGGTACACCTCTGAGGCGGAGGATGACATCCGCTGCGACGCCAGCTGAGAGGACCC 360
 Oy 121 ArgProProArgLeuArgTyrLeuLeuValValSerThrArgGluGluGluLeuSer 140
 Db 361 CGGCGTCCCGGCTCGGTACCTGCTGATGTTTCTACACGAGAAAGAGAGTCTGAGC 420
 Oy 141 GlnAspGluThrValLeuLeuGlyValAspPheProAspSerSerSerProSerCysThr 160
 Db 421 CAGAGTGAAGCGGTCTCTCTGGGCGTGAGATTCTCTGACAGCGAGCTCCGCCACTGACC 480
 Oy 161 LeuGlyLeuValLeuProLeuTrpSerAspThrGlnValTyrLeuAspGlyAspGly 180
 Db 481 CTGGGCTGCTCTTGGCCCTCTGAGTGAACACAGAGTGAATGATGAGAGCGGGGCG 540
 Oy 181 PheSerValThrSerGlyGlnSerArgIlePheLysProIleSerIleGlnThrMet 200
 Db 541 TTCAGCTGAGCTGCTGGTGGCAAGCGGATCTTCAAGCCATCTCCATCCAGACCATG 600

Oy 201 TrpAlaThrLeuGlnValLeuHisGlnAlaCysGlnAlaAlaLeuGlySerGlyLeuVal 220
 Db 601 TCGGCACACTCCAGGTATTGTCACCAAGCATGTAGGAGCTCTTGAAGCGGCTTGTGA 660
 Oy 221 ProGlyGlySerAlaLeuThrTrpAlaSerHisTyrGlnGluArgLeuAsnSerGluGln 240
 Db 661 CCGGCTGGCAGGCGCTTCACTGGGCCAGCCATACAGAGAGACTGTAACCTCCAGACAG 720
 Oy 241 SerCysLeuAsnGluTrpThrAlaMetAlaAspLeuGluSerLeuArgProProSerAla 260
 Db 721 AGCTCCCTCAATAGATGAGCGGCTATGGCCACCGGAGTCTCTGGGCTCCAGCGGCC 780
 Oy 261 GluProGlyGlySerSerGluGlnGlnGlnMetGluGlnAlaIleArgAlaGluLeuTrp 280
 Db 781 GAGCTGGCGGCTCTCAACACAGAGCATGAGACAGGCGATCCGCTGAGCTGTGG 840
 Oy 281 LysValLeuAspValSerAspLeuGluSerValThrSerLysGluIleArgGlnAlaLeu 300
 Db 841 AAGTGTGTGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 860
 Oy 301 GluLeuArgLeuGlyLeuProLeuGlnGlnTyrArgAspPheIleAspAsnGlnMetLeu 320
 Db 900 GAGCTGGCTGGGCTGCCCTCCAGCATGACCTGATGATGATGATGATGATGATGATG 960
 Oy 321 LeuLeuValAlaGlnArg-AspArgAlaSerArgIlePheProHisLeuTyrLeuGly 340
 Db 960 CTGCTGGTGGACAGCGGGGAGCGAGCTCCGCGATCTTCCCTACCTACCTACCTACCT 1019
 Oy 340 rGluTrpAsnAlaAlaAsnLeuGluGlnGlnArgAsnArgValThrHisIleLeuAs 360
 Db 1020 AGAGTGAAGCGCAAACTGGAGAGTGGTGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1079
 Oy 360 mMetAlaArgGluIleAspAsnPheTyrProGluArgPheThrTyrHisAsnValArgLe 380
 Db 1080 CATGCCCCGGGAGATTGACAACTTACCTGAGCGGCTTACCTACCTACCTACCTACCT 1139
 Oy 380 uTrpAspGluGluSerAlaGlnLeuLeuProHisTyrPlyGluThrHisArgPheIleG 400
 Db 1140 CTGGGATGAGAGTGGCCCGAGCTGCTGGCCGCTGGAAGAGAGAGAGAGAGAGAGAG 1199
 Oy 400 uAlaAlaArgAlaGlnGlyThrHisValLeuValHisCysLysMetGlyValSerArg 420
 Db 1200 GGGTGCAGAGACACAGGAGCCAGCCAGTGGTGCCTGACGAACTGGGCTGAGCCCTC 1259
 Oy 420 rAlaAlaThrValLeuAlaTyrAlaMetLysGlnTyrGluCysSerLeuGlnAlaLe 440
 Db 1260 AGCGGCCACAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1319
 Oy 440 uArgHisValGlnGluLeuArgProIleAlaArgProAsnProGlyPheLeuArgGlnLe 460
 Db 1320 GCGCCAGCTGAGAGCTCCGCGCCATGCGCCGCCCAACCTGGCTTCTGGCCAGCT 1379
 Oy 460 uGlnIleTyrGlnGlyIleLeuThrAlaArgThr 471
 Db 1380 GCAGATCTACCAAGGCGATCTGAGCGGCAAGC 1413

RESULT 6 AK094226 2905 bp mRNA linear PRI 15-JUL-2002
 LOCUS AK094226 Homo sapiens cDNA FJ36907 f1s, clone BRACE2003800, weakly similar
 DEFINITION Homo sapiens kinase phosphatase.

ACCESSION AK094226
 VERSION AK094226.1 GI:21753246
 KEYWORDS oligo capping; f1s (full insert sequence).
 SOURCE Homo sapiens cerebellum cDNA to mRNA, clone_11b:BRACE2
 clone:BRACE2003800.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE

1 Tanigami,A., Fujiwara,T., Shibahara,T., Goto,Y., Hiroo,M.,
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 Shimizu,F., Wakebe,H., Ono,T., Hishigaki,H., Watanabe,T., Ozaki,K.,

QY 391 strplysgluthrhisarphellleaglualalaaraglainglythrhisvalleuYa 411
|||||
Db 1380 CTGGAAGGAGACACCGCTTCATTCAGGCTGCAAGAGCAGCAGCAGCAGCTGCTGT 1439
QY 411 THSCTysMetGlyValSerSerSerAlaAlaThrValLeuAlaTyrAlaMetLysG 431
|||||
Db 1440 CCAGCTCAAGATGGGCTGACCGCTCAGCGGCGCACAGTCTGCTATGCCATGGAAGCA 1499
QY 431 nryglucysserleugluGlnAlaLeuArghHisValGlnGlnLeuArpProIleAlaAr 451
|||||
Db 1500 GTACGATGACAGCTGAGAGAGGCGCTGCCGCCACGTCGACAGAGACTCCGGCCATCCGCC 1559
QY 451 gproasnproglyPheLeuArghGlnLeuGlnIleTyrGlnGlyIleLeuThraAla 469
|||||
Db 1560 CCCCAACCTGGGCTCTCGCGCAGCTGCAATCTACAGAGGATCTCCTGACGCGCC 1614

RESULT 7
LOCUS BC028922 2736 bp mRNA linear ROD 07-AUG-2002
DEFINITION Mus musculus, similar to slingshot 3, clone MGC:25738
IMAGE:3987714, mRNA, complete cds.
ACCESSION BC028922
VERSION BC028922.1 GI:20810382
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2736)
REFERENCE 1 Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (01-MAY-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
COMMENT Contact: MGC help desk
Email: cgapds-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Huliyk, S.W., Hale, S.M.,
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAK Plate: 30 Row: 9 Column: 9
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, similarity but not identity to protein.
Location/Qualifiers
1. 2736
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="CZCH IT"
/clone="MGC:25738 IMAGE:3987714"
/tissue_type="Mammary tumor metastasized to lung, Tumor
arose spontaneously from a senescent normal mammary
(clonal) outgrowth infected with the virus MMTV."
/clone_lib="NCT_CGAP_Lu29"
/lab_host="DH10B"
/note="Vector: PCMV-SPORT6"
93..2042
/codon_start=1
/product="similar to slingshot 3"
/protein_id="AAH28922.1"
/db_xref="GI:20810383"

CDS

/translation="MALVTSKSPASGHPVEGPTODRVYRRRIOROSFAVLRG
AVIGLQDGDSENVASEADSEPMEEPSGGEPTEDFTDGGCLQSPDWKQVRHLHLMV
ELLRPDDIRLAALDAPPELRPLRLYLVSNGEELSEEPAILGDEVPSSSCTGL
VPLMSDPOVYLDGDFSVTSGGOSRIFKPVSIOTPMATLOVHLQACVAGGIVP
GGSALAMTYYVOEKINSQOGLNEMMASDSEFPFRPAEFGQASQSEOKMOTIATL
MOVLDTSDDYSTKEIRQALELRIGCLQDQROFTDQMLLLMAQODRASRITPPL
LGSEWMAANLELEOKNRVSHLLNMARELIDNFPPEFTYNNRVWDEESAOLIPHKET
HRFIEDARAOGTRVLVHCKMGVSRSAALVAVAMKQXGMDLEQALIHVOELRPLVRPN
HGFLEOLRTYOGILFASROSHVMEQKVAVSPEEPLAEVSTPLPELPEPGSGGEVM
VMGLEQSEETPEELGLRPRINLRVMSISLSESESTPEAGGLPEVSPDEPL
HPFQSLRAKGGQVRVKKGPWRLKRSQSVVALHSAALVASTRAFPQEGQGOBOSEFG
MSSTPRLRKVMKQASVDSDREDKA"

BASE COUNT 599 a 780 c 589 t
ORIGIN

Alignment Scores:
Pred. No.: 4 84e-133 Length: 2736
Score: 1991.50 Matches: 394
Percent Similarity: 90.43% Conservative: 31
Best Local Similarity: 83.83% Mismatches: 40
Query Match: 81.75% Indels: 5
DB: 10 Gaps: 4

US-09-761-640-4 (1-471) x BC028922 (1-2736)

QY 1 MetaLeuValThrValSerArgSerProProGlySerGlyAlaSerThrProValGly 20
|||||
Db 93 ATGCGCTGCTGACAGTGAATGCTGCGCGCGCAAGTGGCCACATCCAGCTGGGGA 152

QY 21 Pro---TTPaspGlnAlaValGlnArgArgSerArgLeuGlnArgArgGlnSerPheAla 39
|||||
Db 153 CCCAGCAGGACCGAAGTGTAAACGACAGAGCGGCTCCAGCGCAGGAGAGATTTTTCG 212

QY 40 ValLeuArgGlyAlaValLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 59
|||||
Db 213 GTGCTCCGAGAGGCGCTCTCTGAGACTTCAGATGAGAGGAGACAGTATTCATCTGAG 272

QY 60 AlaSerSerGluProThrGluLysAlaProSerGluGlnGlnGlnGlnGlnGlnGlnGln 79
|||||
Db 273 GCTGACTGTGAGCAATGAGAGAACCTCAGGAGAGAGACCCGACGAGCAGACT 332

QY 80 AspPheGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 99
|||||
Db 333 GACAAAGGCAAGACTCCAGACTCCCTGGAACAA---GTCCAAAGAGGAGCAGCTGCAC 389

QY 100 LeuMetValGlnLeuLeuLeuArgProGlnAspPheAlaArgLeuAlaGlnGlnGlnGln 119
|||||
Db 390 CTCATGTGTGAGCTACTGAGCGCGCAGATGATATCCGCTGGCAGCCAGCTGGAGGCA 449

QY 120 ProArgProProArgLeuArgTyrLeuLeuValValSerThrArgGlnGlnGlnGln 139
|||||
Db 450 GCCCGACCCCGCAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 503

QY 140 SerGlnAspGluThrValLeuLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 159
|||||
Db 504 AGT---GAGAGGCACTCTGCTGGGGTGACTTCCCGAAGCAGTTCACAGACAGTGC 560

QY 160 ThrLeuGlyLeuValLeuProLeuThrPheSerAspThrGlnValTyrLeuAspGly 179
|||||
Db 561 ACCCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 620

QY 180 GlyPheSerValThrSerGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 199
|||||
Db 621 GGCTTAGAGTGACATCTGCTGCTCAAAACCGAATCTTCAACCTGCTCTATCCAGAC 680

QY 200 MetTPAlaThrLeuGlnValLeuHisGlnAlaGlyGlnAlaLeuGlnGlnGlnGln 219
|||||
Db 681 ATGTGGCCACACTGAGGTATGACACGAGCATGAGAGTGTAGCAGTGGCTT 740

QY 220 ValProGlyGlySerAlaLeuThrThrPalasSerHisTyrGlnGlnAlaGlnGlnGln 239
|||||
Db 741 GTGCTGTGGGAGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 800

Oy	240	GInSerCySleAsnGluTyrPhrXLaMeTaLaAAspLeuGlusELeArProPseR	259
Db	801	CAGGGTTCCCTCAAGAATGTGAATGCATGTCTGACCGGAGTCCTTGACACCCTAAAT	860
Oy	260	AAlagUpProGIgLYSeSerGelugIngluImetGluGlnAlaIlaArgAlaGluLeu	279
Db	861	GCTGAGCCTGGACACAGGCTTCACAGCAGAAAAGATGAGCACAGGCCATCTGGCTAAATTG	920
Oy	280	TPrLySVallLeuASpValSerASpLeuGlusEValThrSerLySGluIlEAArgAlaGluLeu	299
Db	921	TGGCAGGtGTtGGACACcAGTAGACTGGACAAGCGTCACCTTGTAAGAAGATCCGTGAGGCC	980
Oy	300	LeuGlUleuAlaGrLeuGLYLeuProLeuGInglIntYrArGaSPheThlAASPAnSlmEt	319
Db	.981	CTGGAGCTGCCTCTGGGATGTCTCTCCAGCAGTACCTGCATTTATGTACAAACCATAGT	1040
Oy	320	LeuLeuLeuValAlaGlnArGaSPrgAlaSerAgllIErPeRoHISLeuTyrtLeuGLY	339
Db	1041	TTGGCTGCTATGGCCACCAACAAGACCGGGCCTCTCCATCTTCCACCCTACTTGGGC	1100
Oy	340	SerGIUTrPASnAlaAlaAsnLeuGluGluLeuGlnArGaSPArGVAlThrHisLEu	359
Db	1101	TCTGAGTGGAAATCGTCCCACTGGAGGAACTTCAGAAAAAGAGTAAGTCACTTCTG	1160
Oy	360	ASnMeTalaArgGLUIlaSPAnSndETyrProGIuaRgPheThrTYRHISAnValArg	379
Db	1161	AACATGGCCCCGGAGATTGACAACTTCTCCCTGAGCGCTTCACCTATTACAAATAGCT	1220
Oy	380	LeuTrPaSPGIUGLUSerAlaGlnLeuLeuProHISTrPLYSGLUThrHisArgHeLle	399
Db	1221	GCTGTGGATGAGAAATCGGCACAGCTTGTGCCCTGGAAAGAACACATCGATTCTT	1280
Oy	400	GIuaAlaAlaArgAlaGlnGLYThrHisVallLeuValHIscysLySMeglyValSerArg	419
Db	1281	GAGGATGCCAAGACGACAGGCGACTGGGGTGTAGTCCACTGTAAATAGGTGTCCAGCGCT	1340
Oy	420	SerAlaAlaThrValLeuAlaTyraLaMetLySGIntYrGLUCysSerLeuGluGlnAla	439
Db	1341	TCTGCCCGCCACAGGTACTGAGCCATGTCATGAACAGTATAGCTGGAGCCTGGAGCAAGCC	1400
Oy	440	LeuArGHISVaIGInGLEuArProIIeAlaArgProASnProGILyPheLeuArgGln	459
Db	1401	CTGATCCACGTGCAGGACTCGGCCCATTTGATGACGCCCAACACGAGGCTTCCTGGCCAG	1460
Oy	460	LeuGlnILEtyrGlnGLYlleLeuthrAla	469
Db	1461	CTACGGACCTACACAGCGCATTTCTGACTGCC	1490
RESULT 8	AX398774	2322 bp	DNA linear PAT 27-MAY-2002
LOCUS	AX398774		
DEFINITION	Sequence 10 from Patent WO0220732.		
ACCESSION	AX398774		
VERSION	AX398774.1 GI:21261307		
KEYWORDS	human.		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
REFERENCE	Liou,J.R. 1 Regulation of human map kinase phosphatase-like enzyme Patent: WO 0220732-A 10 14-MAR-2002;		
JOURNAL	Bayer Aktiengesellschaft (DE)		
FEATURES	Location/Qualifiers 1..2322		
SOURCE	/organism="Homo sapiens" /db_xref="taxon:9606"		
BASE COUNT	467 a 751 c 718 g 386 t		
ORIGIN			
Alignment Scores:	7.64e-133	Length:	2322
Pred. No.:			

Score:	1987.50	Matches:	428
Percent Similarity:	86.63%	Conservative:	6
Best Local Similarity:	85.43%	Mismatches:	27
Query Match:	81.59%	Indels:	42
DB:	6	Gaps:	8

US-09-761-640-4 (1-471) x AX398774 (1-2322)

QY	4	ValThrValSerArgSerProPro-----	11
DB	53	GTCACAGTAGAGCGCTTGGCCCCCGGAGGACGCGGGCCCTCCACCGCCCTGTGGGGCCCTG	11
QY	12	-----GlySerGlyAlaSerThrProValGlyProThrPasp-----	23
DB	113	GAATTCCCTGAGAGAGAGAGGGAGGACGCCCTCCCGCCCTCACCCCTGGGG-CTGCCTCT	17
QY	24	-----GlnAlaValGlnArgArgSerArgLeuGlnArgArgGlnSerPheAlaValLeu	41
DB	172	CGGCAGAGACTGGTGTCCACGAGAGAGTGCAGTCCGCGCAAA-----GCTTTG	21
QY	42	Arg-----GlyAlaValLeuGlnGlnAspGlyGlyAspAsnAspAlaIala	58
DB	220	CGGTGCTCTCTGTGGGGGCTGTCTGTGGACTGCAAGATGGAGGGACAATGATGATCGACA	27
QY	59	GluAlaSerSerGluProThrGlyLysAlaProSerGluGluGluLeuHis--GlyAsp	77
DB	280	GAGGCGAGTTCAGGCCAACAAGAAAGCCCCGAGTAGAGAGAGCTCACGGGGCAAC	33
QY	78	GlnThrAsp-PheGly-GlnGlySerGlnSerProGlnLysGlnGluGlnArgGln	97
DB	340	CAGACAGACTTTCGGTGGCAAGGATCCCAAGATCCCAACAAACAGAGAGAGCAGAGCAGC	39
QY	97	IaLeuHisLeuMetValGlnLeuLeuArgProGln-AspAspIleArgLeuAlaIa--	119
DB	400	AACTGCACAC-TCATGAGAGCGCTGCTGAGGCGCCGACGAGTGCATTCGACTTGGAAGCCA	45
QY	116	GlnLeuGluAlaProArgProProArg-LeuArgTyr--LeuLeuValValSerThrAr	134
DB	459	ACGTGAGAGCACCCCGGGGCTCCCGGAGTCCGATACCTTGTTGGTAGTTCTACAG	51
QY	134	Gclu-GlyGluGlyLeuSer-GlnAspGluThrValLeuLeuGly-ValAspPheProAs	15
DB	519	AGAAAGGAGAAAGTGTGAGCCAGGATAGAGAGGCTCCGCGGAGCGGAGTTCCTCGA	57
QY	153	PserSerSerProSerCysThrLeuGlyLeuValLeuProLeuThrPserAspThrGlnVa	173
DB	579	CAGCAGCTCCCCACCTGACCTGCGCTGTCTGTCCCTCTGAGTACACACCGAGT	63
QY	173	IlyrLeuAspGlyAspGlyGlyPheSer-ValThrSerGlyGlnSerArgIlePheLy	193
DB	639	GACTTATATATGAGACGGGGGCTTCAGCGTGCACGTCGTGGGCAAAAGCCGACTTCAA	69
QY	193	SProIleSerIleGlnThrMetTrpAlaThrLeuGlnValLeuHisGlnIaCysGluAl	213
DB	699	GCCCATTCCTCAATCCAGACCAATGTGGGCCACATCCAGTATTGACCAACAGCATGTGAGC	75
QY	213	AlaLeuGlySerGlyLeuValProGlyGlySerAlaLeuThrTrpAlaSerHisTyrGl	233
DB	759	ACCTTAGGACAGCGGCTTGTACCGGGTGGAGTGTCCCTCACCTGGGGCACCTACCA	81
QY	233	nGluArgLeuAsnSerGluGlnSerCysLeuAsnGluTrpThrAlaMetAlaAspLeuGl	253
DB	819	GGAGAGACTGAACCTCCGAACAGAGCTGCTCAATGAGTGAAGCGGTATGGCCGACCTGGA	87
QY	253	USeUeArgProProSerAlaGluProGlyGlySerSerGluGlnGlnMetGluGl	277
DB	879	GTCCTGGGCGCTCCACAGCGCCGAGCCTGGGGGGTCTCAAAACAGGAGCGAGTGGAGCG	93
QY	273	nAlaIleArgAlaGluLeuTrpLysValLeuAspAlaSerAspLeuGlnSerValThrse	293
DB	939	GCGCAATCCGTGTGAGCTGTGAAGATTTGTGATGTGGTACCTCGAGAGATGTGCGCTTC	99
QY	293	IlyGluIleArgGlnAlaLeuGluLeuArgLeuGlyLeuProLeuGlnGlnTrpArgAs	313

```

Db 999 CAGAGAGATCCGCGCAGGCTCTGGAGCTGCGCCGCGGCGCTCCCTCCAGACAGTACCGTGA 1058
Qy 313 pPheIIeaSpaNgInMeIleuLeuValaIaGInaSpaPaIgaIaSerArgIIeph 333
Db 1059 CTTTCATCGCAACACCGATCTCTGCTGTGGCACAGCGGAGCCAGAGCTCCCGCATCTT 1118
Qy 333 eProHIsLeuTyIleuGlySerGIuTrPaSnaIaIaAsnLeuGluGluLeuGlnaArgAS 353
Db 1119 CCCCCACCTCTACCTGGGCTCGAGCTGGAGACCCACCAACCTGGAGAGCTGCAGAGAA 1178
Qy 353 nArgValaIThrHIsIleuAsnMetaIaArgGIuIleaspaSnpheTyIProGIuaRph 373
Db 1179 CAGGCTCACCCGACATCTTG--ATGCCCCGGGAGATTGACAACCTTCACCCGTGACGCTT 1235
Qy 373 eThTyIrhIsaSnVaIaArgLeuTrPaSpGIuGluSerIaGInaLeuLeuProHIsTrPly 393
Db 1236 CACCTACCAACAATGGCCGCTCTGGAGATGAGAGTGGGCCAGCTGCTGCCGCACTGGAA 1295
Qy 393 sGIuThrhIsaRphelIeGluAlaIaArgaIaGInGIyThrhIsaValleuValaIHIsCy 413
Db 1296 GGAGACGACCGCTTCATTTGAGAGCTGCAAGACACAGGACCCACGCTGCTGCCACTG 1355
Qy 413 sLySmetGIyValaSerArgSerIaIaIaThrValleuAlaTyIalAnetIySgInTyrgI 433
Db 1356 CAGATGGGCGCTCAGCCGCTCAGCGGCCACAGCTGCTGCCATGCAATGAAGCAGTACGA 1415
Qy 433 uCySerIeugIuGlnaIaLeuArGHisVaIGInIuLeuArProIleaIaArgProas 453
Db 1416 ATGCAGCTGGAGAGCGCCCTGCGCCACAGTCCAGAGCTCGCGCCATCCGCCGCCCA 1475
Qy 453 nProGIyPheLeuArGIneugInIleTyIeYrGInGIyIleuThraIa 469
Db 1476 CCCTGGCTTCTGCGCAAGCTGCAGATCTACAGGCGCATCTCGACGCGC 1524

RESULT 9
AX405828 2061 bp DNA linear PAT 14-JUN-2002
LOCUS Sequence 243 from Patent WO0222660.
DEFINITION AX405828
ACCESSION AX405828
VERSION AX405828.1 GI:21439095
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1
AUTHORS Tang,Y.T., Liu,C., Zhou,P., Asundi,V., Zhang,J., Zhao,Q.A., Ren,F.,
Xue,A.J., Yang,Y., Wehrman,T. and Drmanac,R.T.
TITLE Novel nucleic acids and polypeptides
JOURNAL Patent: WO 0222660-A 243 21-MAR-2002;
HYSEQ, INC. (US)
FEATURES
source Location/Qualifiers
1..2061
/organism="Homo sapiens"
/db_xref="taxon:9606"
147..1331
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD34935.1"
/db_xref="GI:21439096"
/translation="MALTVSRSPGSGASTPVGPMDAVORRSRLOROSFAVLRKA
VLGLDDGDNDAAEASEPTEKAPSEELHGDTPDGOSGOSFOKQEBQHLHW
OLLRPRODIRLAOLEAPRPRLRYLLVSTRREEGLSODETVLGVDFPSSSPCT
LGLVPLMSDIOVYLDGCGSVTSGGOSRIFKTSIOTMWTLOVLAHQEALIGSG
LVPGSALVTASHTQERLNSQCLNEMADLESRLPSPAEFGSSSEQEMQALR
AELMKVLESTSETSDMPEVFSHESHEPDLQPFOLARTKGQOVDRGPOPALRS
ROSVVTLGGSAVVMRTQAFQEDQGGGEGEPCISSTPRFRKVVAVQASVHSGEHC
EA"
BASE COUNT 415 a 672 c 605 g 369 t
ORIGIN
Alignment Scores:

```

```

Pred. No.: 5,51e-96 Length: 2061
Score: 1470.50 Matches: 290
Percent Similarity: 94.87% Conservative: 6
Best Local Similarity: 92.95% Mismatches: 15
Query Match: 60.37% Indels: 1
DB: 6 Gaps: 1

US-09-761-640-4 (1-471) x AX405828 (1-2061)
Qy 1 MetaIaLeuValaIThrValSerArgSerProProGIySerGIyAlaSerThrProValGIy 20
Db 147 ATGGCCCTGGTCAAGAGTACCGCTTCCGCCCGGAGCGCGCCCTCCACGCGCTGGGG 206
Qy 21 ProTrPaSpGlnaIaValaIaArgArgSerArgLeuGlnaRgaRgInserPheAlaVal 40
Db 207 CCGTGGAGCAGGGGGCTCCAGCAGAGAGTGCAGCTCCAGGAAAGGACAGACTTGGGTG 266
Qy 41 LeuArgGIyAlaValleuGIyLeuGlnaSpGIyIyAspaSnaSpaIaIaIa 60
Db 267 CTCGGTGGGCTGTCTCGGAGCTGCAGAGATGAGAGGAGACAAATGATGATGCAGACAGGC 326
Qy 61 SerSerGIuProThrgIuIySaIaProSerGIuGluLeuHIsGIySpGIuThrAsp 80
Db 327 AGTTTGAGCCACAGAGAGAGGCCCGGAGTGAAGAGAGCTCCACGGGGACAGACAGAC 386
Qy 81 PheGIyGInGIySerGIuSerProGInIySgInGIuGlnaRgInHIsLeu 100
Db 387 TTGGGCAAGAGATCCAGAGTCCCGAGAACAGAGAGACAGAGGAGCAGCTGCACCTC 446
Qy 101 MetValGIaIleuLeuArGProGInaSpaPliLeuGlnaIaIaGInleuGlnaIaPro 120
Db 447 ATGTACACCTGCTGAGCGCCGAGAGTACATCCGCTGCGACCCAGCTGCAGGACACC 506
Qy 121 ArgProProArGIuLeuArGIyIleuLeuValaIaSerThrArgGIuGIyGIuGIyLeuSer 140
Db 507 CGGCTCCCCCGCTCGCTACCTGCTGTAATTTCTACACAGAGAGAGAGAGAGTGTGAC 566
Qy 141 GlnaSpGIuThrValleuLeuGIyAlaSpPheProAspSerSerProSerCysThr 160
Db 567 CAGCATGAGACGCTCTCTGGCGTGTGATTTCCCTGACAGACAGCTCCCGACAGTGCACC 626
Qy 161 LeuGIyLeuValleuProLeuTrPheAspThrGlnaValaIyIleuAspGIyAspGIy 180
Db 627 CTGGGCGTGTCTTCCCTCCCTCGAGTGCACACCCAGGTGTCTTAGAGAGAGAGGGGCG 686
Qy 181 PheSerValThrSerGIyGIyGlnSerArgIlePheIyProIleSerIleGIuThrMet 200
Db 687 TTCACGCTGACCTGTGTGGCAAAAGCCGATCTTCAAGCCCATCTCCATCCACACCATG 746
Qy 201 TrPaIaThrLeuGlnaValleuHIsGlnaIaCysGIuIaIaIaLeuGIySerGIyLeuVal 220
Db 747 TGGGCGACACTCCAGATTTTCACCAACAGCATGTAGAGAGCTTACGACCGGCTTGTA 806
Qy 221 ProGIyGIySerAlaLeuThrTrPaIaSerHIsTyIeGInGIuArgIleuAsnSerGIuGln 240
Db 807 CCGGTGGCAGTCCCTCCTACCTCGGCGCACGACATACAGAGAGACTCACTCCGAACAG 866
Qy 241 SerCysIleuAsnGIuTrPThrAlaMetaIaAspLeuGIuSerLeuArGProProSerAla 260
Db 867 AGCTGCTCAATGAGTGAAGCGCTATGCGGACCTGAGAGTCTGTGGCGCTCCACAGCGCC 926
Qy 261 GIuProGIyGIySerSerGIuGlnGIuGlnIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 280
Db 927 GAGCTGGCGGGCTCTCAGAAACAGAGACATATGAGAGCGATCTCGTGTGAGCTGTGG 986
Qy 281 LysValleuAspVal---SerAspLeuGIuSerValThrSerIySgInIleuArgGlnaIa 299
Db 987 AAAGTGTGGAGCTGGAGACACCACTCAGAGACAGTGCATGCAGAGAGTCTCTCTTCC 1046
Qy 300 LeuGIuLeuArGIleuGIyLeuProLeuGlnGIuTyI 311
Db 1047 CACGAGTCTTCACATGAGAGCTCTTGCAGCCCTTC 1082

```

```

RESULT 10
AK000522      2064 bp      mRNA      linear      PRI 22-FEB-2000
LOCUS
DEFINITION Homo sapiens cDNA FLJ20515 f1s, clone KAT09889.
ACCESSION AK000522
VERSION AK000522.1 GI:7020673
KEYWORDS oligo capping; f1s (full insert sequence).
SOURCE Homo sapiens signet-ring cell carcinoma cell_line:KATO III cDNA to
          mDNA. clone:lib:KAT clone:KAT09889.
ORGANISM Homo sapiens
           Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (sites)
AUTHORS Matanabe,K., Kumagai,A., Itakura,S., Yamazaki,M., Tashiro,H.,
          Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
          Nakamura,Y., Isogai,T. and Sugano,S.
          NEDO human cDNA sequencing project
          Unpublished
          2 (bases 1 to 2064)
          Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
          Shibahara,T., Tanaka,T. and Nakamura,Y.
          Direct Submission
          Submitted (15-FEB-2000) Sumio Sugano, Institute of Medical Science,
          University of Tokyo, Department of Virology, Shirokane-dai, 4-6-1,
          Minato-ku, Tokyo 108-8639, Japan (E-mail:cdna@ims.u-tokyo.ac.jp,
          Tel:81-3-5449-5286, Fax:81-3-5449-5416)
          NEDO human cDNA sequencing project supported by Ministry of
          International Trade and Industry of Japan. cDNA full insert
          sequencing: Research Association for Biotechnology: cDNA library
          construction, 5'- & 3'-end one pass sequencing: Department of
          Virology and Human Genome Center, Institute of Medical Science,
          University of Tokyo (partly supported by Science and Technology
          Agency).
FEATURES
    source
        1..2064
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="KAT09889"
            /cell_line="KATO III"
            /cell_type="signet-ring cell carcinoma"
            /clone_id="KAT"
            /note="cloning vector pME18SFL3"
            147..1331
            /note="unnamed protein product"
            /codon_start=1
            /protein_id="BAA91228.1"
            /db_xref="GI:7020674"
            /translation="MALVTYSRSPSGASTPVGPMDOAVORSLQROSFVLRGA
            VLIQDGGDMDAAEASSEPEKAPSEELHGDQTFGQSSPOKEORHLHLKY
            QLRPDQDILALAEAPRPRLRYLLVSTREGESLODEYILGVDPSSSPSC
            LGLVPLMSDYOYLDGGDFSVYSGGOSRIFRFSIQMTATLOYLHOCFAISG
            LVPFGSALTMASHYORLNSQCLNMTSMADLESLRPSAPRSGSSRDEQOAR
            AELMKVLELSETSSETSDMPVEFSHESHSHEPPLQFPLARTGGQVDVGFQPALKS
            ROSVVTLOGSAVVANRTQAFQEQDQGGQGGEPCLISYTRFRKVVNQAQVHDSGEG
            EA"
BASE COUNT 421 a 671 c 605 g 367 t
ORIGIN
Alignment Scores:
Pred. No.: 5.52e-96 Length: 2064
Score: 1470.50 Matches: 290
Percent Similarity: 94.87% Conservative: 6
Best local Similarity: 92.95% Mismatches: 15
Query Match: 60.37% Indels: 1
DB: 9 Gaps: 1
US-09-761-640-4 (1-471) x AK000522 (1-2064)
QY 1 MetAlaLeuValThrValSerArgSerProGlySerGlyAlaSerThrProValGly 20
    |||||||
Db 147 ATGGCCCTGGGTCACAGTGAAGCCCTTCCGCCCCGGGACAGCGCCCTCCACGCCCGGTGGGG 206
QY 21 ProTrpAspGlnAlaValGlnArgSerArgLeuGlnArgGlnSerPheAlaVal 40

```

```

|||||
Db 207 CCTGGGACCAAGCGGCTCCACAGAGAGTGCATCCAGCGCAAGCGAGCTTGGCGGTG 266
QY 41 LeuArgGlyAlaValLeuGlyLeuGlnAspGlyGlyAspAsnAspAspAlaGlnAla 60
    |||||||
Db 267 CTCCGCGGCGCTGCTCGGAGCTCAGAGATGAGAGGAGCAATGATGATGACAGCAGAGGCC 326
QY 61 SerSerGlyProThrGlyValAlaProSerGlnGlnGlnGlnGlnGlnGlnGlnGln 80
    |||||||
Db 327 AGTTCTGAGCCACAGAGAGAGAGCCCGAGTGAAGAGAGCTCCACGGGAGCCAGACAGAC 386
QY 81 PheGlyGlnGlySerGlnSerProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 100
    |||||||
Db 387 TTGGGCAAGAGATCCACAGAGTCCCGCAGAGCAGGAGGAGCAGGAGCAGCAGCTCAGCTC 446
QY 101 MetValGlnLeuLeuLeuArgProGlnAspAspIleArgLeuAlaGlnLeuGlnAlaPro 120
    |||||||
Db 447 ATGGTACAGCTGTGAGAGCCCGCAGAGATGACATCCGCTGGCAGCCCGCAGTGGAGCACC 506
QY 121 ArgProProArgLeuArgGlyTrpLeuValValSerThrArgGlyGlyGlyLeuSer 140
    |||||||
Db 507 CGGCTCCCGGCTCCGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 566
QY 141 GlnAspGlyThrValLeuLeuGlyValAspPheProAspSerSerSerProSerCysThr 160
    |||||||
Db 567 CAGATGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 626
QY 161 LeuGlyLeuValLeuProLeuThrSerAspThrGlnValTrpLeuAspGlyAspGly 180
    |||||||
Db 627 CTGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 686
QY 181 PheSerValThrSerGlyGlyGlnSerArgIlePheLysProIleSerIleGlnThrMet 200
    |||||||
Db 687 TTCAGCGTGAAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 746
QY 201 TrpAlaThrLeuGlnValLeuHisGlnAlaCysGlnAlaAlaLeuGlySerGlyLeuVal 220
    |||||||
Db 747 TGGCCACACTCCAGAGTATTGTCACAGAGCATGTAGGAGAGCTGTAGGAGAGGCGCTGTGA 806
QY 221 ProGlyGlySerAlaLeuThrTrpAlaSerHisIleGlnGlnGlnGlnGlnGlnGln 240
    |||||||
Db 807 CCGGCTGAGCAGTCCCTCCTGAGGAGCAGCCACATACAGGAGAACTGAACTCCGAAACG 866
QY 241 SerCysLeuAsnGlnIleTrpThrAlaMetAlaAspLeuGlnSerLeuArgProSerAla 260
    |||||||
Db 867 AGCTGCTCATATGATGAGCGGCTATGCGCCAGCTGAGTCTTGGCGCTCCACAGCGCC 926
QY 261 GlnProGlyGlySerSerGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 280
    |||||||
Db 927 GAGCTGGCGGCTCCTCAGAACAGAGAGAGATGAGAGAGGAGGATCCGCTGAGCTGTGG 986
QY 281 LysValLeuAspVal---SerAspLeuGlnSerValThrSerLysGlnIleArgGlnAla 299
    |||||||
Db 987 AAATGTGTGAGCTGAGAGAGCAGCTCAGAGACAGTGCATGCCAGAGGCTCTCTTCC 1046
QY 300 LeuGlnLeuArgLeuGlyLeuProLeuGlnGlnIleThr 311
    |||||||
Db 1047 CACGAGCTTTCACATGAAGAGCTCTGACGCCCTTC 1082

```

```

RESULT 11
AP002776      160903 bp      DNA      linear      HTG 17-JAN-2001
LOCUS
DEFINITION Homo sapiens chromosome 11 clone RP11-126P21 map 11q, WORKING DRAFT
ACCESSION AP002776
VERSION AP002776.2 GI:12246853
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens DNA, clone:RP11-126P21.
ORGANISM Homo sapiens
           Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 160903)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,

```

TITLE
Homo sapiens 160,903 genomic DNA of 11q
Published Only in Database (2000)
REFERENCE
2 (bases 1 to 160903)
AUTHORS
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE
Submitted (10-JUN-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
On Jan 16, 2001 this sequence version replaced gi:9188614.

COMMENT

Center: RIKEN Genomic Sciences Center (GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
Project Information
Center project name: Humdraft1
Center clone name: RP11-126P21
----- Summary Statistics -----
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 149950 bases at least Q40
Consensus quality: 153939 bases at least Q30
Consensus quality: 156144 bases at least Q20
Insert size: 157403; sum-of-contigs
Quality coverage: 8.42x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 36 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

```
1 12626 contig of 12626 bp in length
12727 24539 contig of 11813 bp in length
24640 33605 contig of 8966 bp in length
33706 45529 contig of 11824 bp in length
45630 51226 contig of 5597 bp in length
51327 58567 contig of 7241 bp in length
58668 65488 contig of 6821 bp in length
65589 72315 contig of 6727 bp in length
72416 78297 contig of 5882 bp in length
78398 84113 contig of 5716 bp in length
84214 90536 contig of 6323 bp in length
90637 95051 contig of 4415 bp in length
95152 99283 contig of 4132 bp in length
99384 103658 contig of 4275 bp in length
103759 107823 contig of 4065 bp in length
107924 113135 contig of 5212 bp in length
113136 117634 contig of 4399 bp in length
117635 121241 contig of 3507 bp in length
121242 121341 contig of 100 bp in length
121342 124894 contig of 3553 bp in length
124895 124994 contig of 100 bp in length
124995 128257 contig of 3263 bp in length
128258 128357 contig of 100 bp in length
128358 131408 contig of 3051 bp in length
131409 131508 contig of 100 bp in length
131509 134212 contig of 2704 bp in length
134213 134312 contig of 100 bp in length
134313 136588 contig of 2276 bp in length
136589 136688 contig of 100 bp in length
136689 139497 contig of 2809 bp in length
139498 139597 contig of 100 bp in length
139598 142074 contig of 2477 bp in length
142075 142174 contig of 100 bp in length
142175 143858 contig of 1684 bp in length
143859 143958 contig of 100 bp in length
143959 146242 contig of 2284 bp in length
146243 146342 contig of 100 bp in length
146343 148167 contig of 1825 bp in length
148168 148267 contig of 100 bp in length
148268 150246 contig of 1979 bp in length
150247 150346 contig of 100 bp in length
150347 151876 contig of 1530 bp in length
151877 151976 contig of 100 bp in length
151977 154322 contig of 2246 bp in length
154323 154422 contig of 100 bp in length
154423 155669 contig of 1247 bp in length
155670 155769 contig of 100 bp in length
```

159723 160903 contig of 1181 bp in length.
NOTE: This is a 'working draft' sequence. It currently consists of 36 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```
1 12626 contig of 12626 bp in length
12627 12726: gap of 100 bp
12727 24539: contig of 11813 bp in length
24540 24639: gap of 100 bp
24640 33605: contig of 8966 bp in length
33606 33705: gap of 100 bp
33706 45529: contig of 11824 bp in length
45530 45629: gap of 100 bp
45630 51226: contig of 5597 bp in length
51227 51326: gap of 100 bp
51327 58567: contig of 7241 bp in length
58568 58667: gap of 100 bp
58668 65488: contig of 6821 bp in length
65489 65588: gap of 100 bp
65589 72315: contig of 6727 bp in length
72316 72415: gap of 100 bp
72416 78297: contig of 5882 bp in length
78298 78397: gap of 100 bp
78398 84113: contig of 5716 bp in length
84114 84213: gap of 100 bp
84214 90536: contig of 6323 bp in length
90537 90636: gap of 100 bp
90637 95051: contig of 4415 bp in length
95052 95151: gap of 100 bp
95152 99283: contig of 4132 bp in length
99284 99383: gap of 100 bp
99384 103658: contig of 4275 bp in length
103659 103758: gap of 100 bp
103759 107823: contig of 4065 bp in length
107824 107923: gap of 100 bp
107924 113135: contig of 5212 bp in length
113136 113235: gap of 100 bp
113236 117634: contig of 4399 bp in length
117635 117734: gap of 100 bp
117735 121241: contig of 3507 bp in length
121242 121341: gap of 100 bp
121342 124894: contig of 3553 bp in length
124895 124994: gap of 100 bp
124995 128257: contig of 3263 bp in length
128258 128357: gap of 100 bp
128358 131408: contig of 3051 bp in length
131409 131508: gap of 100 bp
131509 134212: contig of 2704 bp in length
134213 134312: gap of 100 bp
134313 136588: contig of 2276 bp in length
136589 136688: gap of 100 bp
136689 139497: contig of 2809 bp in length
139498 139597: gap of 100 bp
139598 142074: contig of 2477 bp in length
142075 142174: gap of 100 bp
142175 143858: contig of 1684 bp in length
143859 143958: gap of 100 bp
143959 146242: contig of 2284 bp in length
146243 146342: gap of 100 bp
146343 148167: contig of 1825 bp in length
148168 148267: gap of 100 bp
148268 150246: contig of 1979 bp in length
150247 150346: gap of 100 bp
150347 151876: contig of 1530 bp in length
151877 151976: gap of 100 bp
151977 154322: contig of 2246 bp in length
154323 154422: gap of 100 bp
154423 155669: contig of 1247 bp in length
155670 155769: gap of 100 bp
```

* 155770 156932: contig of 1163 bp in length
* 156933 157032: gap of 100 bp
* 157033 158659: contig of 1627 bp in length
* 158660 158759: gap of 100 bp
* 158760 159622: contig of 863 bp in length
* 159623 159722: gap of 100 bp
* 159723 160903: contig of 1181 bp in length.
Location/Qualifiers

FEATURES

source

1. 160903
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q"
/clone="RP11-126p21"
1. 12626
/note="assembly_fragment"
12727. 24539
/note="assembly_fragment"
24640. 33605
/note="assembly_fragment"
33706. 45529
/note="assembly_fragment"
45630. 51226
/note="assembly_fragment clone_end:SP6 vector_side:left"
51327. 58567
/note="assembly_fragment"
58668. 65488
/note="assembly_fragment"
65589. 72315
/note="assembly_fragment"
72416. 78297
/note="assembly_fragment"
78398. 84113
/note="assembly_fragment"
84214. 90536
/note="assembly_fragment"
90637. 95051
/note="assembly_fragment"
95152. 99283
/note="assembly_fragment"
99384. 103658
/note="assembly_fragment"

Alignment Scores:

Pred. No.: 5 12e-89 Length: 160903
Score: 1405.50 Matches: 445
Percent Similarity: 24.94% Conservative: 3
Best Local Similarity: 24.78% Mismatches: 13
Query Match: 57.70% Indels: 1338
Gaps: 13

US-09-761-640-4 (1-471) x AP002776 (1-160903)

QY 10 ProProglySergly-----Alaserthr----- 17
Db 84727 CCAAGGGGCGCGGTAGAAATGCAGATGCTCGACCTCCTAGACCTATCGAC 84786
QY 18 -----ProValGlyProTTrpaspGlnAlaValGln-----ArgArgSerArg 31
Db 84787 GCGTGGGGGCCCATGAACTCCTGTAGACAAAGGCTGCAGTGAATGTGATGCATGCTAGA 84846
QY 32 Leu-----GlnArg 34
Db 84847 CTTTGAGAGCCCTACCTTAAGGGGAATGAGGACCATGACATGTTGCCCTGCCCTGCC 84906
QY 35 ArgGlnSerPheAlaValLeuArgGlyAlaValLeuGlyLeuGlnAspGlyGlyAspAsn 54
Db 84907 AGGACAGACTTTGGGCTCTCGTGGGGCTCTCCTGGGACACGACGAGATGAGAGGACAT 84966
QY 55 AspaPalaAlaGlnAlaSerSerGlnProThrGlnLysAlaProSerGlnGlnLysLeu 74
Db 84967 GATGATGACAGAGGCGCATTTCTGAGCCAAACAGAGAGGCCCGCCAGTGAAGAGAGCTC 85026

QY 75 HisGlyAspGlnThrAspPheGlyGlnGlySerGlnSerProGlnLysGlnGln 94
Db 85027 CACGGAGACACAGACACTTGGGGGAGAGATCCAGATCCCAAGAGAGAGAGAGAG 85086
QY 95 ArgGlnHisLeuHisLeuMetValGlnLeuLeuArgProGlnAspPheArgLeu--- 113
Db 85087 AGGACGACCTGCACCTCATGTGTACAGTGTGAGGCGCGAGATGACATCCGCTGGT 85146
QY 113 ----- 113
Db 85147 AGGCCCATGTGAGCTCCGGGGGTGGGGGAAGAGACACGCTGAGAGAGATGGAG 85206
QY 113 ----- 113
Db 85207 CCTGTGTATGTGTCTGCTTGGGAGACATGTGTGAGGCGAGGGGCGAGCTGGGAGT 85266
QY 113 ----- 113
Db 85267 GTTGGGGTTACCATATCCAGTGACCTGCTGTTTCAGAGAGAGATGTTGTTTTTG 85326
QY 113 ----- 113
Db 85327 TTTGTTTTTTTTTGGAGATGAGTCTGCTGTGTCAGGCTGAGAGTGGCGTG 85386
QY 113 ----- 113
Db 85387 ATCTCGCTCACTGCAAGTTCACCTCCGGGTTTCACACCATTTCTCGCTCAGACCTCC 85446
QY 113 ----- 113
Db 85447 CGAGAGCTGGGACTACAGCACCCGCCACTATGCTGTGTAATTTTTTTGTATTTT 85506
QY 113 ----- 113
Db 85507 GTAGAGATGCGGTTTACCGCTTAGACGAGGTGTCTCAATCTCTGACTCGTATTC 85566
QY 113 ----- 113
Db 85567 GCGTGGCTCGGGCTCCCTAAGTGTGGGATTACAGGCTGAGCCACCGCGCGGCTGG 85626
QY 113 ----- 113
Db 85627 GCAGAGATATTTTGACAGGCGGAGTCAATGCTGAGCTCTGTGAAGGAAGAACAG 85686
QY 113 ----- 113
Db 85687 GGAGGTTTGTCTGTGTGAGCCAAATAGGCCAAATTTAGTCCAAAGGGGTGGAAATTCAG 85746
QY 113 ----- 113
Db 85747 AGGCATGTTTCTCCTCAACACAGTAGAAGACTTCTCTGATGATGTGAGCTGCCCTAATG 85806
QY 113 ----- 113
Db 85807 GACCACTCACAGGACGCTTGTGTGACACACTGTGGGGAGGTGTGTAGAAGGAC 85866
QY 113 ----- 113
Db 85867 CAGGACAGAGGAGAGAGACTTCAGCCCCCAGAGAGAGAGATCCAGTGAAGCTGGAG 85926
QY 113 ----- 113
Db 85927 CGCTCATGCTCAAACTACTTGTGGAGACAGCTCGCCGCTCTCAGCAAGACGTGGCAACAA 85986
QY 113 ----- 113
Db 85987 GATGAGTGAATCTTTGTAAACTGAATTTACAACAGAGATCATCCCTATGTGAG 86046
QY 113 ----- 113
Db 86047 ATTTTGTCTTAATATTTCTTTGTGATTAACATTTCTCAATGAATGATAGATTTGC 86106
QY 113 ----- 113

OY 301 |uLeuArgLeuGlyLeuProLeuGlnGlnIleTyrArgAspPheIleAspAsnIleLeu 321
 DB 88326 AGCTCGGCTGGGGCTCCCGCTCCAGCAGTACCGTACTTCAACACACGATGCTGC 88385
 OY 321 euLeuValAlaGlnArgAspArgAlaSerArgIlePheProHisLeuTyrLeu 338
 DB 88386 TGTCTGTGGACAGCGGAGCGAGCGCTCCCGCATCTTCCCGACCTTACTGTGAGCT 88445
 OY 338 ----- 338
 DB 88446 TCAGCCAGCGCTGGGCCATGGGGACCGCTGCAGCTGTGAGGGCGGACAGGAGAGCAG 88505
 OY 339 ----- glySerGlnTyrPasnAla 344
 DB 88506 AGCTGGAGAGAGAGAGAAATGTGCTGTTCCTCTCTCCAGGGCTCAGATGGAAACGA 88565
 OY 345 AlaAsnLeuGlnGluLeuGlnArgAsnArgVal ----- 355
 DB 88566 GCAAACTGTGAGAGAGCTGCAGAGAGAAAGTGA - GGGCTATGAGCCCTCGGGCCACCCAC 88624
 OY 355 ----- 355
 DB 88625 CCCATCTTCCCTTCTCTCGGCCCTCCCGCATTTGGGTGTGAGCAGACTTCAAAACCCCTG 88684
 OY 355 ----- 355
 DB 88685 GACCACTCTCAGCAGCTGCTAGCTGCTTCTTAACTGTGTCTGCGGGCTGTTCCTGCT 88744
 OY 355 ----- 355
 DB 88745 GTGGGCTCCAGGTGGGGACAGAGAGACCTGCTGTGGACCGCCCGCCACTCTCTCCCCC 88804
 OY 355 ----- 355
 DB 88805 ATCCACTGTGAAAACAAGACAGAAAGAGGCGCTCAGCCAGCCCAAGACAGAGAGCA 88864
 OY 355 ----- 355
 DB 88865 GCAGGCGACTGCTGTGTAAGTCTTGGACAGCAAGAAAAAGGCTCTCTTGAATGCGCC 88924
 OY 355 ----- 355
 DB 88925 TGTGGGCCAGCTACTTGGAGAGGTGAGGACAGAGAGATGCTTGAAGCCTGAGATTTAG 88984
 OY 355 ----- 355
 DB 88985 GCGCAGTGTAGCCGTGATCAGCGCACACTGCAGCCTGGGCAACAGAGAGAGACCTG 89044
 OY 355 ----- 355
 DB 89045 TCTTAATAAATAAGAAAAAGAGAGAAAAAGCCTTTTCTCCACCTTGCCTGTCT 89104
 OY 355 ----- 355
 DB 89105 CAGGAGAAGAAGACTGCCCTTCTCCCGTGGGACCTGGCTGCTGTGACAGGTAC 89164
 OY 355 ----- 355
 DB 89165 CTGTATCTGCCCCACCATGAGGCTTGTGGAGACTGCTAGCCCCCTGCACACCATGCTGC 89224
 OY 355 ----- 355
 DB 89225 AGACCCACCCACTCTCAGCTTAGCTCAAAAAGCTGTCTTAAGTCAATTTCTGAGATAAC 89284
 OY 355 ----- 355
 DB 89285 TGAAGGGCTGAGATTGCAAGTTGGCCAGCTGTCTGACACAGATGGGGAACAAGCCAGC 89344
 OY 355 ----- 355
 DB 89345 AGGCAAGATGATGTGTTAAGTGTGACAGCCAGGTGACAGCTGGGTCACTTCTCTCCCA 89404

OY 355 ----- 355
 DB 89405 CTGTCACTGTGCTCCCATCTGACTGAACACTCAGAGCTAAAGGAGGCCAGGGGCTGC 89464
 OY 355 ----- 355
 DB 89465 CAGTCCGATTGCCAGTGGGCTGTACTCTGCCCACATGAGCCTGGCTTGGGCCCTC 89524
 OY 356 ----- ThrHisIleLeuAsnMetAlaArgGluIleAspAsnPheTyrProGlu 371
 DB 89525 TGGACAGGGTCACCACATCTTGAACATGGCCGGGAGATTGACAACTTACCTTGAG 89584
 OY 372 ArgPheThrTyrHisAsnValArgLeuTyrPaspGlnGlnSerAlaGlnLeuProHis 391
 DB 89585 CCCTTACCCTACACAAATGTGGGCTCTGGAGTAGAGATGAGCGCCACACTGCTGCGGCAC 89644
 OY 392 TrpTyrGlnTyrHisArgPheIleGluAlaAla ----- 402
 DB 89645 TGAAGAGAGACGACCGCTTCAATGAGCTGC - AAGTCTGCTCTGCTGCTTGTGCT 89703
 OY 402 ----- 402
 DB 89704 GCCCAGTGTGGCTCAAGCGGCTCGGGTCCCTCTCTGCCACATCTGCATGAGCT 89763
 OY 402 ----- 402
 DB 89764 GTGCATCTCTCCAGCCCTCAGTGTCTCTCCCTCTCTACTCTCAGTTGTTTGA 89823
 OY 402 ----- 402
 DB 89824 GCCTTCACTGCTCCCTCCCTCCAGAGGCGGAGATCTTCATGTTGAGGGAGGCTGCTG 89883
 OY 403 ----- ArgAlaGlnGlyThrHisValLeuValHisCysIleMetGly 416
 DB 89884 AACCTGGCGCTGCTCCACAGACAGGACAGGACCGACCGCTGTGCTCCATGCAAGATGGCG 89943
 OY 417 ValSerArgSerAlaAlaThrValLeuAlaTyrAlaMetLysGlnTyrGluCysSerLeu 436
 DB 89944 GTACGCCGCTCAGCGGCCACAGATGCTGGCCTATGCGCATGAAAGAGTACGATGACGCTG 90003
 OY 437 GluGlnAlaLeuArgHisValGlnGlnLeuArgProIleAlaIleArgProAsnProGlyPhe 456
 DB 90004 GAGCAGGCCCTGCGCCAGGTGAGAGCTCCGGCCATGCGCCGCCCAACCTGCTTC 90063
 OY 457 LeuArgGlnLeuGlnIleTyrGlnGlnIleLeuThrAlaArg 470
 DB 90064 CTGCGCAGCTGCAGATCTACAGGAGCATCTGAGCGGACAG 90105
 RESULT 12
 LOCUS AP001885 171566 bp DNA linear HTG 13-JAN-2001
 DEFINITION Homo sapiens chromosome 11 clone RP11-157K17 map 11q, WORKING DRAFT
 SEQUENCE, 24 unordered pieces.
 ACCESSION AP001885
 VERSION AP001885.3 GI:12225033
 KEYWORDS HTG; HTGS; PHASE1; HTGS; DRAFT.
 SOURCE Homo sapiens DNA, clone:RP11-157K17.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 171566)
 AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 TITLE Homo sapiens 171,566 genomic DNA of 11q
 JOURNAL Published Only in Database (2000)
 REFERENCE 2 (bases 1 to 171566)
 AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 TITLE Direct Submission
 JOURNAL Submitted (24-Apr-2000) Masahira Hattori, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
 1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 (E-mail:hattori@gs.c.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,

COMMENT

Tel:81-45-503-9111, Fax:81-45-503-9170)
On Jan 13, 2001 this sequence version replaced gi:8117536.

----- Genome Center

Center: RIKEN Genomic Sciences Center(GSC)

Center code: RIKEN

Web site: <http://hgp.gsc.riken.go.jp/>

Contact: hatoriegsc.riken.go.jp

----- Project Information

Center project name: Humdraft11

Center clone name: RP11-157K17

----- Summary Statistics

Sequencing vector: PCR products; 100% of reads

Chemistry: Dye-terminator ET-amersham; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 164093 bases at least Q40

Consensus quality: 166550 bases at least Q30

Consensus quality: 167995 bases at least Q20

Insert size: 169266; sum-of-contigs

Quality coverage: 8.92x in Q20 bases; sum-of-contigs

----- NOTE: This is a 'working draft' sequence. It currently consists of 24 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

```

1      15114 contig of 15114 bp in length
15215 27885 contig of 12671 bp in length
27986 42419 contig of 14434 bp in length
42520 55000 contig of 12481 bp in length
55101 68523 contig of 13423 bp in length
68624 79107 contig of 10484 bp in length
79208 89207 contig of 10000 bp in length
89308 95794 contig of 6487 bp in length
95895 103725 contig of 7831 bp in length
103826 112105 contig of 8280 bp in length
112206 119213 contig of 7008 bp in length
119314 124281 contig of 4968 bp in length
124382 130027 contig of 5646 bp in length
130128 135012 contig of 4885 bp in length
135113 138400 contig of 3288 bp in length
138501 144921 contig of 6421 bp in length
144921 150480 contig of 5459 bp in length
150481 150580 contig of 100 bp
150581 153837 contig of 3257 bp in length
153838 153937 contig of 100 bp
153938 156948 contig of 3011 bp in length
156949 157048 contig of 100 bp
157049 161384 contig of 4336 bp in length
161385 161484 contig of 100 bp
161485 164413 contig of 2929 bp in length
164414 164513 contig of 100 bp
164514 167396 contig of 2883 bp in length
167397 167496 contig of 100 bp
167497 169991 contig of 2495 bp in length
169992 170091 contig of 100 bp
170092 171566 contig of 1475 bp in length.

* NOTE: This is a 'working draft' sequence. It currently
* consists of 24 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1      15114: contig of 15114 bp in length
15115 15214: gap of 100 bp
15215 27885: contig of 12671 bp in length
27886 27985: gap of 100 bp
27986 42419: contig of 14434 bp in length
42420 42519: gap of 100 bp
42520 55000: contig of 12481 bp in length
55001 55100: gap of 100 bp
55101 68523: contig of 13423 bp in length
68524 68623: gap of 100 bp
68624 79107: contig of 10484 bp in length
79108 79207: gap of 100 bp
79208 89207: contig of 10000 bp in length
89208 89307: gap of 100 bp

```

FEATURES

source

```

*      89308      95794: contig of 6487 bp in length
*      95795 95894: gap of 100 bp
*      95895 103725: contig of 7831 bp in length
*      103726 103825: gap of 100 bp
*      103826 112105: contig of 8280 bp in length
*      112106 112205: gap of 100 bp
*      112206 119213: contig of 7008 bp in length
*      119214 119313: gap of 100 bp
*      119314 124281: contig of 4968 bp in length
*      124282 124381: gap of 100 bp
*      124382 130027: contig of 5646 bp in length
*      130028 130127: gap of 100 bp
*      130128 135012: contig of 4885 bp in length
*      135013 135112: gap of 100 bp
*      135113 138400: contig of 3288 bp in length
*      138401 138500: gap of 100 bp
*      138501 144921: contig of 6421 bp in length
*      144922 145021: gap of 100 bp
*      145022 150480: contig of 5459 bp in length
*      150481 150580: gap of 100 bp
*      150581 153837: contig of 3257 bp in length
*      153838 153937: gap of 100 bp
*      153938 156948: contig of 3011 bp in length
*      156949 157048: gap of 100 bp
*      157049 161384: contig of 4336 bp in length
*      161385 161484: gap of 100 bp
*      161485 164413: contig of 2929 bp in length
*      164414 164513: gap of 100 bp
*      164514 167396: contig of 2883 bp in length
*      167397 167496: gap of 100 bp
*      167497 169991: contig of 2495 bp in length
*      169992 170091: gap of 100 bp
*      170092 171566: contig of 1475 bp in length.

FEATURES
source
1. 171566
   /organism="Homo sapiens"
   /db_xref="taxon:9606"
   /chromosome="11"
   /map="11q"
   /clone="RP11-157K17"
1. 15114
   /note="assembly-fragment clone_end:T7 vector_side:right"
15215..27885
   /note="assembly-fragment"
27986..42419
   /note="assembly-fragment"
42520..55000
   /note="assembly-fragment"
55101..68523
   /note="assembly-fragment"
68624..79107
   /note="assembly-fragment"
79208..89207
   /note="assembly-fragment"
89308..95794
   /note="assembly-fragment"
95895..103725
   /note="assembly-fragment"
103826..112105
   /note="assembly-fragment"
112206..119213
   /note="assembly-fragment"
119314..124281
   /note="assembly-fragment"
124382..130027
   /note="assembly-fragment"
130128..135012
   /note="assembly-fragment"
135113..138400
   /note="assembly-fragment clone_end:sp6 vector_side:left"
138501..144921
   /note="assembly-fragment"
145022..150480
   /note="assembly-fragment"

```

```

/misc_feature      /note="assembly_fragment"
150581. .153837
/note="assembly_fragment"
/misc_feature      153938.156948
/note="assembly_fragment"
/misc_feature      157049. .161384
/note="assembly_fragment"
/misc_feature      161485. .164413
/note="assembly_fragment"
/misc_feature      164514.167396
/note="assembly_fragment"
/misc_feature      167497. .169991
/note="assembly_fragment"
/misc_feature      170092. .171566
/note="assembly_fragment"
BASE COUNT      45656 a 39885 c 40004 g 43721 t 2300 others
ORIGIN
Alignment Scores:
Pred. No.:      5.54e-89      Length:      171566
Score:          1405.50      Matches:      445
Percent Similarity: 24.94%      Conservative: 3
Best Local Similarity: 24.78%      Mismatches: 13
Query Match:    57.70%      Indels:      1338
DB:              2      Gaps:      13
US-09-761-640-4 (1-471) x AP001885 (1-171566)
QY 10 ProProGlySerGly-----AlaSerThr----- 17
Db 201 CCAAGGGGCCCGGCTTAGAAATGACAGATGCTCAGGCTCCACTCTAGACCTATCGACC 260
QY 18 -----ProValGlyProThrPaspGlnAlaValGln-----ArgArgSerArg 31
Db 261 GCGTGGGGGCCCATGAAATCCGTGTAGACAAAGGCTGCAGGTGAATGTGATGCATGAGA 320
QY 32 Leu----- 34
Db 321 CTTTGAGAGCCCTACCCCTAGGGGAATGAGAGCCATGACAGTTGCCCATCTGCCCTGCC 380
QY 35 ArgGlnSerPheAlaValLeuArgGlyAlaValLeuGlyLeuGlnAspGlyGlyAspAsn 54
Db 381 AGGAGAGACTTGGCGGTCTCCGTGGGGCTGTCTCGGAGCTGCAAGATGAGAGGGACAAAT 440
QY 55 AspaPalaIaGlnAlaSerSerGlnProThrGluLysAlaProSerGlnGlnGlnLeu 74
Db 441 GATGATGACAGCAGAGGCGCATGTCGAGCCAAACAGAGAGGCCCGCAGTGAAGAGAGCTC 500
QY 75 HisGlyAspGlnThrAspPheGlyGlnGlySerGlnSerProGlnIlyGlnGlnGln 94
Db 501 CACGGGGGACACAGACAGACTTCGGCCAAAGATCCCAAGATCCCAAGACAGAGAGAGCAG 560
QY 95 ArgGlnHisLeuHisLeuMetValGlnLeuLeuArgProGlnAspAspIleArgLeu--- 113
Db 561 AGGCAGACCTGCACCTCATGTGTACAGCTGTAGAGCGCCAGAGATGACATCCGCTGGTG 620
QY 113 ----- 113
Db 621 AGGGCCCATGTGAGCTCCGGGGGTGGGGGAAGACACACCTGAGGGCAGAAATGAG 680
QY 113 ----- 113
Db 681 CCCTGTGTATGTCTGTCTGTTGGGAGCAATGTTGAGAGGGCAGGGGCGAGCTGGGAGT 740
QY 113 ----- 113
Db 741 GTTGGGGTTACCATATCCACAGTGACCTGTCTTTCAGGACAGAGATGTTTGTGTTTGG 800
QY 113 ----- 113
Db 801 TTTGTTTTTTTTTTTGTAGATGAGTCTGCCTGTGTTGCCAGGCTGGAGTGCAGTGGCCTG 860
QY 113 ----- 113

```

```

Db 861 ATCTGGCTCACTGCAGAGTTCACACTCCCGGGTTACACACATTCCTGCTCAGACCTCC 920
QY 113 ----- 113
Db 921 CGAGAGCTGGAGCTACAGACGCCGCCACTATGCTGGCTAAATTTTTTTTGTATTTTGA 980
QY 113 ----- 113
Db 981 GTAGAGATGCGGTTTACACCGTGTGTTAGCCAGGGTGTCTCAATCTCTGACCTCGTATCC 1040
QY 113 ----- 113
Db 1041 GCCTGCCCTCGGCTCCCTAAGTGTGGGATTACAGGCGTGAGCCACGCCGCCGCCCTGG 1100
QY 113 ----- 113
Db 1101 GCAGAGATATTTTGCACAGGGCGAGGTGATGCTGAGGTCTCTGGAAAGAAAGACAGG 1160
QY 113 ----- 113
Db 1161 GGAGGTTTGTCTCTGTGTGAGCCCAATAGGGCCCAATTAGGTTCCAAAGGGGTGGGAATTCAG 1220
QY 113 ----- 113
Db 1221 AGGCATGTTCTCCTCAACACAGTAGAAGACTTTCCTGATGATGTAGCTGTCCCTAATG 1280
QY 113 ----- 113
Db 1281 GACCATGTGCACAGGACAGGCTTGTGTGACACACACTGGGGGAGGGTGTGTAGAGGACC 1340
QY 113 ----- 113
Db 1341 CAGGACAGGGGAGAGGAGACTTCCAGCCCAAGAGAGAAAGGATGCCAGTGTGAGCTGAG 1400
QY 113 ----- 113
Db 1401 CGCTCATCTCAATATCTTGTGTGAGACAGCTGCCGCTCTCAGCAAGACGTGGCAACAA 1460
QY 113 ----- 113
Db 1461 GATCAGTGAATCTTGTGTAAACTGAATTTACAAACAGAGTATCCTCATCTGTGAG 1520
QY 113 ----- 113
Db 1521 ATTTGTCTTAATATTTCTTGTGTGATAAACATTTCTCAATGAAATGATGACATTTGC 1580
QY 113 ----- 113
Db 1581 AAGTGACAGGGTTAAAAAACAACAAAAAAGAACTGGCGGGGGGGGTGGCTCATGCC 1640
QY 113 ----- 113
Db 1641 TGTATCCCACTACTTGGGAGCGGAGCGGCGAGATCACAGAGTGCAGAGATCGAGAC 1700
QY 113 ----- 113
Db 1701 CATCTGCTTAACACAGGTGAACCCCGTGTCTACTAAAAACACAAAAAATTAGCCAGGCG 1760
QY 113 ----- 113
Db 1761 TGTGGCGGGTGCCTGTATGCTCCAGCTACTCGGAGGCTGAGGACAGGAATGGGCTGA 1820
QY 113 ----- 113
Db 1821 CCCGGAGATGAGAGTTGCACTGAGCTGAGATTGGCCACTGCCAGCTGGGTGAC 1880
QY 113 ----- 113
Db 1881 AGAGCAAGACTCCATTTGCCAAAAAAGAAAAAAGAAAAAAGAACTCTTTCTTGTGCGCACAGT 1940
QY 113 ----- 113

```

Db	1941	GGCTACGCGCTGTAAATCCAGCACTTTTGGGAGGCTGAGAGGAGAGATGTGGCTAGTCCA	2000
Qy	113	-----	113
Db	2001	GGAGTTCAGACACAGCCACAGAAACATAGGAGATCTCATCTCAACAAAAACTTAATAA	2060
Qy	113	-----	113
Db	2061	ATTATCTGGGCATGTGTGGCATGTGGGCTGATGTGCCAGCAACTTGGGAGGCTAAGGTGGGA	2120
Qy	113	-----	113
Db	2121	GGATCATGTAGAGCCTGGGAGTTGGGGGCTACAGTAGAGCTGTGATGTGCCACTGTACTCC	2180
Qy	113	-----	113
Db	2181	AGCCTGGGCGACAGAGTAGAGACCCATCTCAAAACAAAAACAAACAACTCCTCTCC	2240
Qy	113	-----	113
Db	2241	AGCAGAAAACAAAAAATGTGCACACATATGTCACTGCCATTCCTCCGCTTTTTCACACAGT	2300
Qy	113	-----	113
Db	2301	CCCTCTTACCCACTGATTTTACATCTGGGAGAGGGGGGATGGAGAGGAGAGACAGG	2360
Qy	113	-----	113
Db	2361	GTCAGACAGTGTCTGGGCTGGGGGTCTAGTGGGGAGACAGGCTCTTGGGGCCACTGT	2420
Qy	114	-----A1A1A1G1n1eug1uA1P1o1r1P1r1o1P1r1o1A1r1G1	2480
Db	2421	GACCGTGGGTCTCTCATCTCCCGCCAGCGACGCCAGCTGGAGGCAACCCCGCTCCCGG	2480
Qy	125	LeuArGTyR1euleuVal1SerThrArGg1uG1yG1uG1yLeuSerG1n1aSPeG1uThr	144
Db	2481	CTCCGCTACTGCTGCTGATGTTTACACAGAGAGAGAGAGTCTGAGCCAGGATGAGCG	2540
Qy	145	ValLeu1eug1yVal1aSPheProaSPSer	154
Db	2541	GTCCTCTGGGCGGATTTCCCTGACAGCAGGTTTCAGACGAGAGAGAAAGAGGGGC	2600
Qy	155	-----SerSerProse	158
Db	2601	AAAGAGGTGAGGGACAGGGGGGCAATGCTTCCCTCTGTCCCTGCCACACTCCCCAG	2660
Qy	158	rCySTh1eug1yLeuVal1leuPro1eu1rPse1aSPTh1rG1n1a1y1yR1euaSPG1yAs	178
Db	2661	CTGACCCCTGGGCTGTGTGCTGCCCTCTGGAGTACACCCAGGTGTACTTATGATGAGA	2720
Qy	178	pG1y-----	179
Db	2721	CGG-GTAAAGCAATGGCAACTGAGAGTGGGGGCTGGAGGSGTGAATTAACTAGAGTGAAGG	2779
Qy	179	-----	179
Db	2780	ATGTGACGGGGCTTGGGACACAGGTATACAGTAACCTTATAGCTGGAGCTTGGGGCTG	2839
Qy	179	-----	179
Db	2840	CTCCCACTCCACGTACAGTTCACCGTGATCCTGAGCAAGGACCTGACTCCTGGGTCTC	2899
Qy	179	-----	179
Db	2900	GGCTTCCCGTATCCAGCAAAAGATGGTTCTGTGCGAGAGCCTGGAGTCTGGCTCA	2959
Qy	180	-----G1yPheSerVal1Th1SerG1yG1n1SerArg1lePh	192
Db	2960	CCTGCGCTGGTCTCTCGACGAGGGCTTACGCGTACAGCTGTGGTGGCAACCGGATCTT	3019
Qy	192	e1ySProl1eSer1leG1n1Th1rMet1rP-----	201
Db	3020	CAACCCATCTCCATCTCCAGACCATGTGTGTAGAGACAGACACTGCTGGACTACAGCCAGC	3079

[illegible]

Oy	355	-----	355	-----	355
Db	4159	GACCACCCTCAGCAGCTGCTGACTTGCTTCTTAACCTGTCTGCGGGCTGTGGCCCTGGT	4218	4159	GACCACCCTCAGCAGCTGCTGACTTGCTTCTTAACCTGTCTGCGGGCTGTGGCCCTGGT
Oy	355	-----	355	-----	355
Db	4219	GTGGGCTCCAGGTTGGGGACAGGAGACTGTGTGGCAGGCCCCGCCACTCTCTCCCCC	4278	4219	GTGGGCTCCAGGTTGGGGACAGGAGACTGTGTGGCAGGCCCCGCCACTCTCTCCCCC
Oy	355	-----	355	-----	355
Db	4279	ATCCACACTGTGAACACAGGACAGAAACAAGGGCCTCAGCCAGCCCAAGACGAGAACA	4338	4279	ATCCACACTGTGAACACAGGACAGAAACAAGGGCCTCAGCCAGCCCAAGACGAGAACA
Oy	355	-----	355	-----	355
Db	4339	GCACGCATACCTGCTGTAACTGCTTGGACACAGAAAAAGGCTCTCTTGAATGCGCC	4398	4339	GCACGCATACCTGCTGTAACTGCTTGGACACAGAAAAAGGCTCTCTTGAATGCGCC
Oy	355	-----	355	-----	355
Db	4399	TGTGGGCCACGACTTCTTGGGAGGCTGAGCAGAGAGATGGCTTGAGCCCTGAGATTTAG	4458	4399	TGTGGGCCACGACTTCTTGGGAGGCTGAGCAGAGAGATGGCTTGAGCCCTGAGATTTAG
Oy	355	-----	355	-----	355
Db	4459	GCCCGATGAGCCGTGATCACGCCACTGCATCCAGCTGGGCAACAGAGAGACCTG	4518	4459	GCCCGATGAGCCGTGATCACGCCACTGCATCCAGCTGGGCAACAGAGAGACCTG
Oy	355	-----	355	-----	355
Db	4519	TCCTTAAAAATAAGAAAAAGAAAGAAAAAGCCCTTTCTCCACCTTGCCCTGT	4578	4519	TCCTTAAAAATAAGAAAAAGAAAGAAAAAGCCCTTTCTCCACCTTGCCCTGT
Oy	355	-----	355	-----	355
Db	4579	CAGGGAAGAAAGAACTGCCCTTCTCCCCGTTGGGAGACTGTGGCTGCTGTACAGGTAC	4638	4579	CAGGGAAGAAAGAACTGCCCTTCTCCCCGTTGGGAGACTGTGGCTGCTGTACAGGTAC
Oy	355	-----	355	-----	355
Db	4639	CTGTATCTGCCCCACCATGGGCTTCTGGGACCTGTGTAGCCCTTGCCACCATGCTGC	4698	4639	CTGTATCTGCCCCACCATGGGCTTCTGGGACCTGTGTAGCCCTTGCCACCATGCTGC
Oy	355	-----	355	-----	355
Db	4699	AGACCACCCACTCTCAGCTTAAGCTCAAAAAGCTGTTCTTAACGATTTCTGAGATAAC	4758	4699	AGACCACCCACTCTCAGCTTAAGCTCAAAAAGCTGTTCTTAACGATTTCTGAGATAAC
Oy	355	-----	355	-----	355
Db	4759	TGAAGGGCTGGAGTTGCAGTTGGCCCAAGCTGTCTGGACCAAGTTGGGAAACAAGCCAGC	4818	4759	TGAAGGGCTGGAGTTGCAGTTGGCCCAAGCTGTCTGGACCAAGTTGGGAAACAAGCCAGC
Oy	355	-----	355	-----	355
Db	4819	AGGCGAAGATGATTGTTGTTAAGTTCGACAGCAGGTGACAGCTGGGTCACCTTCTCCCA	4878	4819	AGGCGAAGATGATTGTTGTTAAGTTCGACAGCAGGTGACAGCTGGGTCACCTTCTCCCA
Oy	355	-----	355	-----	355
Db	4879	CTGTACCTGCTGCCCTCCATCTGACTTGAACATCAGAGCTAAAGCGAGGCCAGGGGCTGC	4938	4879	CTGTACCTGCTGCCCTCCATCTGACTTGAACATCAGAGCTAAAGCGAGGCCAGGGGCTGC
Oy	355	-----	355	-----	355
Db	4939	CAGTCCGATTGCCAGTGGGCGCTGGTAACCTTGCCCACTCAGCCCTTGGGCGCTC	4998	4939	CAGTCCGATTGCCAGTGGGCGCTGGTAACCTTGCCCACTCAGCCCTTGGGCGCTC
Oy	356	-----	371	-----	371
Db	4999	TGGGACAGGGTACACCCACATCTTGAAACATGGCCGGGAGATTGACAACTTCAACCTAG	5058	4999	TGGGACAGGGTACACCCACATCTTGAAACATGGCCGGGAGATTGACAACTTCAACCTAG
Oy	372	ArgPheThrTyrrHisAsnValArgLeuTTPasPglUserAlaGlnLeuLeuProHis	391	372	ArgPheThrTyrrHisAsnValArgLeuTTPasPglUserAlaGlnLeuLeuProHis
Db	5059	CGCTTCACTACCAACCAATGTGGCCCTCTGGATGAGAGACTCGGCCAGCTGCTGCCGAC	5118	5059	CGCTTCACTACCAACCAATGTGGCCCTCTGGATGAGAGACTCGGCCAGCTGCTGCCGAC
Oy	392	TrpLysGlnPheHisArgPheIleGluAlaIaIa-----	402	392	TrpLysGlnPheHisArgPheIleGluAlaIaIa-----
Db	5119	TGGAGGAGAGCAGCCGCTTATTGAGCTGC-AAAGTGGTCTCCTGGCTTGTCTGCTCT	5177	5119	TGGAGGAGAGCAGCCGCTTATTGAGCTGC-AAAGTGGTCTCCTGGCTTGTCTGCTCT
Oy	402	-----	402	-----	402
Db	5178	GCCCCAGTGGGCTTCAAGGGGCGCTCCGGTGCCCTCTCTGCCCCACTGTGCCATCAGCT	5237	5178	GCCCCAGTGGGCTTCAAGGGGCGCTCCGGTGCCCTCTCTGCCCCACTGTGCCATCAGCT
Oy	402	-----	402	-----	402

Db	5238	GTGCATTCTTCACGACCCTAGTGTCTCCTCCCTCTCTACTCTCAGTTGGTTGA	5297
Oy	402	-----	402
Db	5298	GCCCTTCACCTGCCCTCCCTCGACGAGCGGGATCTTACGTGTGAGGGAGGTGCTG	5357
Oy	403	-----	403
Db	5358	AACCTGGCCTGCTTCACAGACACAGGACACCCACGTGCTGTCACCTGCAGATGGGC	5417
Oy	417	VAlSerArgSerAlaAlaThrValLeuAlaTrpAlaMetLysGlnTyrglucYsserLeu	436
Db	5418	GTACAGCCCTCAGCGGCCACAGTGCTGTGCCTCATGCAATGAAAGCAATGACGCTG	5477
Oy	437	GlUglnAlaLeuAlaArgHisValGlnLeuArgProIleAlaArgProAsnProGlyPhe	456
Db	5478	GAGCAGGCCCTGCGCACGTGTGAGAGGCTCGGGCCATCGCCCCAACCCCTTGCTTC	5537
Oy	457	LeuArgGlnLeuGlnIleTyrglnGlnIleuThrAlaArg	470
Db	5538	CTGCGCCAGCTGCAGATCTACGAGGCAATCTCTGACGCCCAGG	5579
RESULT	13		
AB072356	2280 bp	mRNA	linear PRI 29-JAN-2002
LOCUS	Homo sapiens mRNA for hSSH-1S, complete cds.		
DEFINITION	AB072356		
ACCESSION	AB072356.1	GI:18376660	
VERSION			
KEYWORDS			
SOURCE	Homo sapiens CDNA to mRNA.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
TITLE	1		
JOURNAL	Ntwa,R., Nagata-Ohashi,K., Takelchi,M., Mizuno,K. and Uemura,T.		
MEDLINE	Control of actin reorganization by Slingshot, a family of phosphatases that dephosphorylate ADF/cofilin		
CITATION	Cell 108 (2), 233-246 (2002)		
PMID	21822082		
DOI	2 (bases 1 to 2280)		
FEATURES	Location/Qualifiers		
source	1..2280		
gene	/organism="Homo sapiens"		
cds	/db_xref="taxon:9606"		
	1..2280		
	/gene="hSSH-1S"		
	1..2079		
	/gene="hSSH-1S"		
	/standard_name="slingshot-1S"		
	/note="dual specific phosphatase"		
	/codon_start=1		
	/product="hSSH-1S"		
	/protein_id="BAB4115.1"		
	/db_xref="GI:18376661"		
	/translation="MALVTIORSPPTSAASSASNSLEAGSEFDRKLTLSISPFMVKCALFLQSSSPQGRSLDHPKKHAGDLPOLQWLNILRCEDRIKLAVRESANARAKVAVVVYSGRQDTRENLGVDSKSRSCTIGMTVLRLMSDTIKILDGGGFSVSTAGRHLIRPVSVQAMWSALQLVLRACEVARHNYPGGVALIMATVYESCISSQSIENEMAMDLESTRDPSPALFVDKRTGEGERTRFKALKRISTMSQDLENVSKIEIRNELKEOMCNCLKEJEFIDENMLLIIGOMKRPGLIPDHLYLGSENNASNLEELOGSGEVDLYLVNTREINDNEFGPLPAVHNIRRVYDEDETDLAHMNEYHYHINKAKRHSHCLVHKMGVSASASTVIAYAMKRWGMPELKAYNVYKKORSTTRPAGMRQDSETEGLIDS KQRNKLMKROUTSDSLQQPVDDPGPDGFLETPETRESPQLDDAADQGCLGPLDCFRNRKLDPLPSPEDETGSLVTHLEDERALEEAAPAEVHRPARQPOOGSGLCERK		


```
/standard_name="slingshot-1L"
/note="dual specific phosphatase"
/codon_start=1
/product="hSSH-1L"
/db_xref="GI:18376659"
/translation="MALVTLQSRPTPSAASSASNSLEAGSEEDRKLTNLSSEFPM
VKGAALFLQOGSSPOGSRLOHPKHHAGDLPQHLQVMI NLRCDRIKLAVLSAMA
DRVRYVMVYVSSGSDTEENTLLGVDSKESKSTQGMVILRLMSDRIHLDGGGGS
VSTAGRMHFKRPVSVOAMMSALQVHLKACSEVARRHNPFGVALIMATVYSCISSE
SCIENMANMODLESTRPDSPALFYDKPTBEGRTRLIKAKRISIMSDOLENTYKEL
RNELEKQMNCKLKEFIDNEMLLIGQMDKPSLITDHLAIHNEVYHFINAKRHHSCVLA
VDILNVTREIDNEPGLFAHNI RYDEETDILA HNEVYHFINAKRHHSCVLA
CKMGSVASTVIAVYAMKEFGMPLEKAYNYKQKRSITRPMAGMRQDLESEGLDAS
KORHNKLMROOTDSLOQVDDPAGPDFLEPTDGPESOLPELDADQAGLPPLP
CCFRRLSDPLLPSEDETSGLVHLEDEBERALIEEAPPAVHRRAPROOGSGLCCK
DYKRLKPEFGSKGSGSLLOVEYTEREBGLAGWQOLPTDLDNLNENLNNNSKR
SCFMGEMDDAIFGLINKVKPSYKSCADCMPTASGAPASREDEDPAATCTOPAF
LPHTSSPVVHLASRSRVPEKPAEPPEPEPLPAPSRADYTGSGAGALPEPASL
LEPSETPKVHLKSLILKNSHCDKNPSTEVIIEEESPKMDKMPARDRLLEFENSE
KPTNSYLMQHOESIIOLKAGLVKHTKELERLSYADAPAPSRDGPASRLBASIP
EESODPAALHELIGPLVMPGSGKDESEAPASLEGGSLKPPPEFYRLDHTSXSSP
FLKTICTPTSSSSMSNLTRSSSDSIHVRKRGVLYKQRIQLEIETRLIAGLVSSP
LKRSHTLAKLSLTFSTEDUSSEADPSTVADSOJTTSESEFLHEPGTDPDAATSK
PSCKPAPEMNKSPSMWSK"
```

```
BASE COUNT      973 a      1100 c      994 g      749 t      1 others
ORIGIN
Alignment Scores:
Pred. No.:      3,996-70      Length:      3817
Score:          1113.00      Matches:      232
Percent Similarity: 65.14%      Conservative: 80
Best Local Similarity: 48.43%      Mismatches: 127
Query Match:      45.69%      Indels:      40
DB:              9      Gaps:      8
```

US-09-761-640-4 (1-471) x AB072355 (1-3817)

```
Oy 1 MetalaleuValThrValSerArgSerProProGlySerGlyAlaSerThrProValGly 20
Db 1 ATGGCCCTTGAGACCCCTGACGCGCTGCCACGCCACGCCGCGCTCCGCCGCCAC 60
Oy 21 ProTrpAsp-----GlnAlaValGlnArgSerArgLeuGlnArgGln 36
Db 61 AACACGAGGTGGAGCTGCGACGACGACAGATCGAATAATTAACTCAGCTTAAGTAG 120
Oy 37 SerPheAlaValLeuArgGlyAlaValLeuGlyLeuGlnAspGlyGlyAspAsnAsp 56
Db 121 AGCTTTTCATGCTGAAGGCGACGCTCTCTTACCAAGGGA-----165
Oy 57 AlaAlaGluAlaSerSerGluProThrGluLysAlaProSerGluGluGluLeuHisGly 76
Db 166 -----ACGAGCCCTCAA-----177
Oy 77 AspGlnThrAspPheGlyGlnGlySer---GlnSerProGlnLysGlnGlu---GlnGln 94
Db 178 -----GGCCAGGCGAGCTCTCAGACGACCCGCCAACAGACGAGGTGATCTCG 222
Oy 95 ArgGlnHisLeuHisLeuMetValGlnLeuLeuArgProGlnAspAspIleArgLeuAla 114
Db 223 CCTCAACATCTTCAGTCAATCAACCTCTGCGTCCGAAAGAGATCAAGATCAAGCTGGCA 282
Oy 115 AlaGlnLeuGluAlaProArgProArgLeuArgTyrLeuLeuValVal-----Ser 132
Db 283 GTGCCCTCGAGAGCGCTGGCGGACGCGGCTCCGCTACATGCTGTGTGACGACGACC 342
Oy 133 ThrArgGluGlyGlyGlyLeuSerGlnAspGlnThrValLeuLeuGlyValAspPhePro 152
Db 343 GGGCGCCAG-----GACACCGAGGAGATATCTTCGCGAGGTGAGTTCCTTCC 390
Oy 153 AspSerSerSerProSerCysThrLeuGlyLeuValLeuProLeuTrpSerAspThrGln 172
Db 391 AGTAAGAAAGTAAAGCTGCACCATTTGGAGTTCGCGAGTGTGAGGACGACGAGAA 450
```

```
Oy 173 ValTyrLeuAspGlyAspGlyIypPheSerValThrSerGlyGlyGlnSerArgIlePhe 192
Db 451 ATCCACCTTGATGAGGATGTGTGGTTACCGCCGACACAGCAGAGATGACATATT 510
Oy 193 LysProIleSerIleGlnThrMetTrpAlaThrLeuGlnValLeuHisGlnAlaCysGlu 212
Db 511 AAGCTGTGTCTGTCCAGGCCATGTGTGTGCTGCGCTGACAGGTCTTCAACAAGCTGGCAA 570
Oy 213 AlaAlaLeuGlySerGlyLeuValProGlyGlySerAlaLeuThrTrpAlaSerHisTyr 232
Db 571 GTGGCCGAGGACACACTACTTCCCGGGGTGTAGCTCTATCTGCGCTACACTAT 630
Oy 233 GlnGluAlaGlyLeuAsnSerGluGlnSerCysLeuAsnGluTrpThrAlaMetAlaAspLeu 252
Db 631 GAGACTGATCATCGCTCCGACGAGCTGCATCAACAGCTGAGAACCCATGAGACCTG 690
Oy 253 GluSerLeuArgPro-----ProSerAlaGluProGlyGlySerSerGluGlnGln 270
Db 691 GAGTCTACGCGCGCCGACTCCCGCGCTATTGTGACAAAGCCCACTGAAGGGAAG 750
Oy 271 MetGluGlnAlaIleArgAlaGluLeuTrpLysValIleAspValSerAspLeuIser 290
Db 751 ACCGAGCGCTCATCAAGCCAAAGCTCCGAAACATCATGATGAGCGAGGATCTGAAAAT 810
Oy 291 ValThrSerLysGluIleArgGlnAlaLeuGluLeuArgGlyLeuProLeuGlnGln 310
Db 811 GTGACTTCCAAAGACATTCGTAATGATTAAGAAACAGATCAATTGACTTGAAGGA 870
Oy 311 TyrArgAspPheIleAspAsnGlnMetLeuLeuValAlaGlnArgAspArgAlaSer 330
Db 871 CTCAGAGATTTATGACAAATGAGATGACTTATCTTGCGGACGATGAGCAAGCCCTCC 930
Oy 331 ArgIlePheProHisLeuTyrIleuGlySerGluTrpAsnAlaIleAsnLeuGluLeu 350
Db 931 CTATCTTCGATCATCTTATCTCGCTGATGAATGATCAATCTGAGGAGAACTG 990
Oy 351 GlnArgAsnArgValThrHisIleLeuAsnMetAlaArgGluIleAspAsnPheTyrPro 370
Db 991 CAGGCTCAGGGGTGATTTACTATTAAATGTTACCAGAAATGCAATATTTTCTCT 1050
Oy 371 GluArgPheThrTyrHisAsnValArgLeuTrpAspGluGlnSerAlaGlnLeuPro 390
Db 1051 GCGTATTGTCATATCATATCAATCCGAGCTCATGATGAAGACACAGACTCTCTGCC 1110
Oy 391 HisTrpLysGluThrHisArgPheIleGluAlaAlaArgAlaGlnGlyThrHisValLeu 410
Db 1111 CACTGGAATGAAGCGTATCATTTTATAACAAAGCGAAGAGAACCATTCCAAGTGCCTG 1170
Oy 411 ValHisCysLysMetGlyValSerArgSerAlaIleThrValLeuAlaTyrAlaMetLys 430
Db 1171 GTGCATTGCCAANAATGGCGGTGAGTGCCTCGGCTCCACAGTCATAGCTATGCAATGANG 1230
Oy 431 GlnTyrGluCysSerLeuGlnGlnAlaLeuArgHisValGlnGluLeuArgProIleAla 450
Db 1231 GAATTCGGCTGCGCTCGGAAAGACATATACATATGTAACAGACAGCGAGCTACG 1290
Oy 451 ArgProAsnProGlyPheLeuArgGlnLeuGlnIleTyrGlnGlyIleLeuThrAla 469
Db 1291 CGCCCAACGCGGCGCTTATAGAGGAGCTGTGATGATGAAGCATCTTGATGCA 1347
RESULT 15
AX180874
LOCUS AX180874 6374 bp DNA linear PAT 06-AUG-2001
DEFINITION Sequence 1 from Patent WO0146394.
ACCESSION AX180874
VERSION AX180874.1 GI:15132702
KEYWORDS
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 (bases 1 to 6374)
```

AUTHORS Plovman,G.D., Martinez,R., Whyte,D., Manning,G., Sudarsanam,S.,
Hill,R.J. and Flanagan,P.
TITLE Mammalian protein phosphatases
JOURNAL Patent: WO 0146394-A 1 28-JUN-2001;
Sugen, Inc. (US)
FEATURES Location/Qualifiers
source 1. .6374
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1604 a 1674 c 1659 g 1437 t
ORIGIN
Alignment Scores:
Pred. No.: 7,49e-70 Length: 6374
Score: 1113.00 Matches: 232
Percent Similarity: 65.14% Conservative: 80
Best Local Similarity: 48.43% Mismatches: 127
Query Match: 45.69% Indels: 40
DB: Gaps: 8
US-09-761-640-4 (1-471) x AXI80874 (1-6374)
QY 1 MetAlaLeuValThrValSerArgSerProGlySerGlyAlaSerThrProValGly 20
|||||
DB 34 ATGGCCCTGGTACCCCTGCAGCCGCCACGCCGCCCTCTCTCCGCCAGC 93
QY 21 ProTyrASP-----GlnAlaValGlnArgArgSerArgLeuGlnArgGln 36
|||
DB 94 AACAGCAGTGGAGGCTGCACGAAAGATCGAAATTAACCTCAGCTTAAGTAG 153
QY 37 SerPheAlaValLeuArgGlyAlaValLeuGlnAspGlyAspAspAsp 56
|||||
DB 154 AGCTTTTCATGTGAAAGCGCAGCCCTCTTACAAAGGGA----- 198
QY 57 AlaAlaGlnAlaSerSerGluProThrGlnAlaProSerGluGluLeuHisGly 76
|||
DB 199 -----AGCACCCTCA----- 210
QY 77 AspGlnThrAspPheGlyGlnGlySer---GlnSerProGlnInsGlnIu---GluGln 94
|||||
DB 211 -----GGCCGCGGAGGCTCTTCACACCCCCACAGCATGAGTGTCTG 255
QY 95 ArgGlnHisLeuHisLeuMetValGlnLeuLeuArgProGlnAspPheArgLeuAla 114
|||||
DB 256 CCTCAACATCTTCAGGATGATCAACTCTCTCGTTGCAAGACAGAAATCAAGCTGCA 315
QY 115 AlaGlnLeuGlnAlaProArgProProArgLeuArgTyrLeuLeuVal-----Ser 132
|||||
DB 316 GTGCGCTGGAGAGCCCTGGCGGACCGGTCGGTACATGGTGTGTACAGCAGC 375
QY 133 ThrArgGlnGlyGlnGlyLeuSerGlnAspGluThrValLeuLeuGlyValAspPhePro 152
|||||
DB 376 GGGCGGCAG-----GACACCGAGAGAAATATCTCTGCGAGAGGACTTTTCC 423
QY 153 AspSerSerSerProSerCysThrLeuGlyLeuValLeuProLeuThrSerAspThrGln 172
|||
DB 424 AGTAAGAAAGTAAGAGCTGCACATTTGGATGTCTCGACTGTGAGCGACAGAAA 483
QY 173 ValTyrLeuAspGlyAspGlyGlyPheSerValThrSerGlyGlyGlnSerArgIlePhe 192
|||||
DB 484 ATCCACCTTGATGAGATGGTGTTCAGGTACGACACAGAAAGATGATGATATTTT 543
QY 193 LysProIleSerIleGlnThrMetTyrPheAlaThrLeuGlnAlaLeuHisGlnAlaCysGlu 212
|||||
DB 544 AAGCGTGTGTCTCCAGGCGCATGTGTCTGCCCTCAGGTCTTACAAAGGCTGCGAA 603
QY 213 AlaAlaLeuGlySerGlyLeuValProGlyGlySerAlaLeuThrTyrPheSerHisTyr 232
|||
DB 604 GTGGCCCGAGGACAACTACTTCCCGGGGGTGTAGCTCTCATCTGGCTACTACTAT 663
QY 233 GlnGlnArgLeuAsnSerGlnInsSerCysLeuAsnGluTyrPheAlaMetAlaAspLeu 252
|||||
DB 664 GAAGAGTGCATGCTCGCAGAGAGCTGCATCAACAGAGTGAACGCATGCGAGAGCTG 723

QY 253 GluSerLeuArgPro-----ProSerAlaGlnProGlySerSerGluGlnGln 270
|||||
DB 724 GAGTCTACGGGCCCCGACCTCCGCCCTATTGTGGACAAAGCCCTGAGGGAAGG 783
QY 271 MetGlnAlaIleArgAlaGlnLeuTyrPheValLeuAspValSerAspLeuGlnSer 290
|||||
DB 784 ACCGAGGCGCTCATCAAGCCCAAGCTCCGAGACATCATGTGAGCCAGGATCTAGAAAT 843
QY 291 ValTyrSerLysGlnIleArgGlnAlaLeuGlnLeuArgLeuGlyLeuProLeuGlnGln 310
|||||
DB 844 GTGACTTCCAAAGAGATTCGAATGAAATTAGAGAAACAGATGAAATTGTAACTTGAAGAA 903
QY 311 TyrArgAspPheIleAspAsnGlnMetLeuLeuValAlaGlnArgAspArgAlaSer 330
|||||
DB 904 CTCAGGAATTTTATAGCAATGAGATGCTTATCTTGGGACGATGAGCAAGCCCTCC 963
QY 331 ArgIlePheProHisLeuTyrLeuGlySerGluTyrPheAsnAlaAlaAsnLeuGlnGlu 350
|||||
DB 964 CTATATCTTCATCATCTTTATCTCGGCTCGAATGGAATCATCCATCTCGAGAGAACTG 1023
QY 351 GlnArgAsnArgValThrHisIleLeuAsnMetAlaArgGluIleAspAsnPheTyrPro 370
|||||
DB 1024 CAGGGCTCAGAGGGTGTGATTTACATTTAAATGTTACAGAGAAATCGATATTTTTCCT 1083
QY 371 GlnArgPheThrTyrHisAsnValArgLeuTyrPheGluSerGluSerAlaGlnLeuPro 390
|||||
DB 1084 GGCTTATTTCATATATATTAACATCCAGTCTACGATGAAGAGCACAGACCTCTCGCC 1143
QY 391 HisTyrLysGlnThrHisArgPheIleGlnAlaAlaArgAlaGlnGlyThrHisValLeu 410
|||||
DB 1144 CACTGGAATGAACGATATCATTTTATTAACAAAGCAAGAGAACCATTCGAAGTCCCTG 1203
QY 411 ValHisCysLysMetGlyValSerArgSerAlaAlaThrValLeuAlaTyrAlaMetLys 430
|||||
DB 1204 GTGCATTTGCAAAATGGCGGAGTCCCTCGCTCCACAGCTCATACCTTAATGAATGAAG 1263
QY 431 GlnTyrGlnCysSerLeuGlnGlnAlaLeuArgHisValGlnGlnLeuArgProIleAla 450
|||||
DB 1264 GAATTTGGCGTGGCTGTGAAAGAGCATATTAATATGTAAGCAAGAGCGACATCAGC 1323
QY 451 ArgProAsnProGlyPheLeuArgGlnLeuGlnIleTyrGlnGlnIleLeuThrAla 469
|||||
DB 1324 CGCCCAACGCGGCTTTATGAGGAGAGCTGTGTGAGTATGAAGGATCTTGATGCA 1380

Search completed: April 12, 2003, 09:51:32
Job time : 7629 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 11, 2003, 00:12:38 ; Search time 887 Seconds

(without alignments)
1195.819 Million cell updates/sec

Title: US-09-761-640-4
Perfect score: 2436
Sequence: 1 MALTVSRSPGSGASTPVG.....PNEGFLKQLQIYQGLTART 471

Scoring table:

BLOSUM62			
Xgapop 10.0 ,	Xgapext 0.5		
Ygapop 10.0 ,	Ygapext 0.5		
Fgapop 6.0 ,	Fgapext 7.0		
Delop 6.0 ,	Delext 7.0		

Searched: 2185239 segs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODBL-frame+ p2n model -DEV-xlh
-Q=/cgnr_1/USPRO.spool/us09761640/runat_08042003_141434_20345/app-query.fasta.1.647
-DB=N_Geneseq_101002 -QFWT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pcr -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=us09761640@cgn.1.1.450 @runat_08042003_141434_20345 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT LONGLOG -DEV.TIMEOUT=120
-WARP.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq_101002: *
1: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT: *
2: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT: *
3: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT: *
4: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT: *
5: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT: *
6: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT: *
7: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT: *
8: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT: *
9: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT: *
10: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT: *
11: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT: *
12: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT: *
13: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT: *
14: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT: *
15: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT: *
16: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT: *
17: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT: *
18: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT: *
19: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT: *
20: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT: *
21: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT: *
22: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001.DAT: *
23: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT: *
24: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2436	100.0	2618	24	AAD36063	Murine dual-specific
2	2436	100.0	2704	24	ABO73249	Human MAP kinase p
3	2436	100.0	2704	24	ABO73252	Human MAP kinase p
4	2426	99.6	2718	24	AAD36061	Human dual-specific
5	2426	99.6	2781	22	AAF30485	Human protein phos
6	2426	99.6	2852	24	ABO73250	Human MAP kinase p
7	2061.5	84.6	2540	24	ABO73251	Human MAP kinase p
8	1987.5	81.6	2322	24	ABL08005	Human MAP kinase p
9	1470.5	60.4	2061	24	ABN59832	Human MAP kinase p
10	1113	45.7	6374	22	AAD09493	Human SGP006 phosph
11	1076.5	44.2	2260	22	AAD09493	Human SGP001 phosph
12	1045	42.9	1711	22	AAD12966	Human dual-specific
13	1004.5	41.2	1771	22	AAH22966	Human dual-specific
14	980	40.2	1949	22	AAD12965	Human dual-specific
15	918.5	37.7	4467	23	ABL10739	Drosophila melanog
16	852	35.0	3488	24	ABL57466	Human protein phos
17	793	32.6	1026	22	AAF63578	Human phosphatase
18	760.5	31.2	1755	22	AAH14722	Human MAP kinase p
19	760.5	31.2	1755	24	ABL40801	Human MAP kinase p
20	733	30.1	8002	23	ABL10738	Drosophila melanog
21	714	29.3	717	22	AAH98183	Human ESP-derived
22	689	28.3	599	24	ABL40803	Human MAP kinase p
23	687	28.2	409	24	ABL40802	Human MAP kinase p
24	648	26.6	1348	22	AAH99712	Human protein enco
25	591.5	24.3	426	24	ABL40800	Human MAP kinase p
26	589	24.2	1052	22	AAD12967	Human dual-specific
27	568.5	23.3	571	22	AAH07057	Human cDNA clone (
28	539	22.1	969	22	AA59843	Human novel cytol
29	438	18.0	1450	22	AA541387	cDNA encoding nove
30	438	18.0	1450	22	AA543834	cDNA encoding nove
31	413	17.0	494	24	ABL40804	Human MAP kinase p
32	401	16.5	447	21	AACT5065	Human ORFX ORF620
33	357	14.7	951	21	AA233341	Human secreted pro
34	269	11.0	828	22	AAF63567	Murine phosphatase
35	256	10.5	901	22	AAF63576	Human phosphatase
36	256	10.5	1087	21	AA635094	Human cellular pro
37	256	10.5	1292	23	AA581801	Human 18221 cDNA.
38	256	10.5	1574	24	AA923605	Human protein phos
39	235.5	9.7	1337	24	AAK99409	DNA of APP related
40	235.5	9.7	1830	21	AA664262	Human dual-specific
41	235.5	9.7	2192	21	AAAF7672	DNA encoding a hum
42	231.5	9.5	687	22	AAF29608	Murine DSP-3 varia
43	231.5	9.5	1067	22	AAF63565	Murine phosphatase
44	227.5	9.3	2050	22	AAF63577	Human phosphatase
45	227.5	9.3	2118	22	AAF30479	Human protein phos

ALIGNMENTS

RESULT 1
AAD36063

AAD36063 standard; cDNA; 2618 BP.

AC AAD36063;

DT 09-AUG-2002 (first entry)

DE Murine dual-specificity phosphatase 15 (DSP-15) cDNA.

XX Murine; dual-specificity phosphatase 15; DSP15; anti-allergic; cytosolic;

XX Immunosuppressive; MAP; mitogen activated protein kinase; cancer; enzyme;

XX Signal transduction; cell proliferation; Duchenne muscular dystrophy;

XX cell cycle abnormality; graft-versus-host disease; autoimmune disease;

XX metabolic disease; allergy; screening; gene; ss.

XX Mus musculus.

OS

XX

FH	Key	Location/Qualifiers
FT	CDS	35..1450
FT		/*tag= a
FT		/product= "Murine DSP-15 protein"
XX		
PN	MO200224740-AZ.	
XX		
PD	28-MAR-2002.	
XX		
PF	19-SEP-2001; 2001WO-US29406.	
PP		
PR	19-SEP-2000; 2000US-233833P-	
PR	18-SEP-2001; 2001US-0955732.	
XX		
PA	(CEPT-) CEPTYR INC.	
PI	Luche RM, Wei B;	
DR	WPI: 2002-394127/42.	
DR	P-PDB; AAE22733.	
XX		
PT	New dual-specificity phosphatase 15 polypeptide and polynucleotides,	
PT	useful for treating e.g. Duchenne muscular dystrophy, cancer,	
PT	grft-versus-host disease, autoimmune diseases, allergies, metabolic	
PT	diseases -	
XX		
PS	Claim 56; Fig 4; 91pp: English.	
CC	The invention relates to a new isolated dual-specificity phosphatase 15	
CC	(DSP-15) polypeptide which retains the ability to dephosphorylate an	
CC	activated MAP (mitogen activated protein) kinase. DSPs are phosphatases	
CC	that dephosphorylate both phosphotyrosine and phosphothreonine/serine	
CC	residues. DSP-15 polypeptides may be used to identify agents that	
CC	modulate DSP-15 activity, where such agents may inhibit or enhance signal	
CC	transduction via a MAP-kinase cascade, leading to cell proliferation. DSP	
CC	polypeptides, modulating agents, and/or polynucleotides encoding the	
CC	polypeptides may be used to modulate DSP-15 activity in a patient, and to	
CC	ameliorate disorders such as Duchenne muscular dystrophy, cancer, graft-	
CC	versus-host disease, autoimmune diseases, allergies, metabolic diseases,	
CC	abnormal cell growth, abnormal cell proliferation and cell cycle	
CC	abnormalities. DSP-15 alternate form polypeptides are useful in screening	
CC	assays for modulators of enzyme activity and/or substrate binding. The	
CC	present sequence is murine DSP-15 cDNA.	
XX		
SQ	Sequence 2618 BP; 538 A; 857 C; 769 G; 454 T; 0 other:	
	Alignment Scores:	
	Pred. No.: 4,51e-195 Length: 2618	
	Score: 2436.00 Matches: 471	
	Percent Similarity: 100.00% Conservative: 0	
	Best Local Similarity: 100.00% Mismatches: 0	
	Query Match: 100.00% Indels: 0	
	DB: 24 Gaps: 0	
	US-09-761-640-4 (1-471) x AAD36063 (1-2618)	
OY	1 MetAlaLeuValThrValSerArgSerProGlySerGlyAlaSerThrProValGly 20	
Db	35 AMGGCCCTGGTCAACATGTAGCCCGTTGCGCCCGGGCAGCGCGCTCCACGCCCGTGCGG 94	
OY	21 ProTPASgLIAlAValAlGlnArgArgSerArgLeuGlnArgArgGlnSerPhAlaVal 40	
Db	95 CCCTGGGACCAGCGCGGTCCAGCAGAGAGTCGACTCCAGCGAAGCGCAGAGCTTTGCGCGTG 154	
OY	41 LeuArgGlyAlAValLeuGlyLeuGlnAspGlyGlyVAspAsnAsPARAdAlAGlAlA 60	
Db	155 CTCGCTGGGGCTCTCCTGGGACTGCGAGATGGAGGGGACAATGATGTGTCAGCGAGAGGCC 214	
OY	61 SerSetGIuPrOThrGInLyuAsAlAProSerGInGInGluIleuHisGlyAspGInThAsp 80	
Db	215 AGTTCTGAGCACACAGAAGAGCCCCGAGAGTAGAGGAGAGCTCCAGCGGGGACCAACAGAC 274	
OY	81 PheGlyInGlySerGlnSerProGlnLysGlnGluGlnAlnArgGlnHisLeu 100	

Db	275	TTCCGGCAGAGATCCCAAGACTCCCCAGAACAGGAGACAGAGCCAGACTGCACCTC	334
Qy	101	MetValGlnLeuLeuArgProGlnAspAspLLeaArgLeuAlaAlaGlnLeuGlnAlaPro	120
Db	335	ATGGACAGAGCTGCTGAGCGCCGACGAGATGACATCCGCTCGGACGCCAGCTGGAGGACCC	394
Qy	121	ArgProProAlaGLeuAlaGlyTyrLeuLeuValValSerThrArgGlnGlyGlnLeuSer	140
Db	395	CGGGCTCCCGGGCTCCCTCCACTCCCTGGTAGTTCTTACACGGAAAGAAAGATGCTGAGC	454
Qy	141	GlnAspGlnThrValLeuLeuGlnGlyValAspPheProAspSerSerSerProserCysThr	160
Db	455	CAGGATGAGACGGTCTCTCTGGCGGTGGATTTCTCTGAACGACGCTCCCGCAGCTGCACC	514
Qy	161	LeuGlyLeuValLeuProLeuThrPheAspThrGlnValTyrLeuAspGlyAspGlyGly	180
Db	515	CTGGGCGCGGTCTGGCCGCTGGAGTGAAGACACCAGGTGACTAGATGAGACGGGGGC	574
Qy	181	PheSerValThrSerGlyGlnSerArgLLeuPheLysProLLeuSerLLeuThrMet	200
Db	575	TTTCACCGTGAAGCTGTGGTGGCAAGCGGGATTTTCACACCCATCTCCATCCAGACATG	634
Qy	201	TrpAlaThrLeuGlnValLeuHisGlnAlaCysGlnAlaAlaLeuGlySerGlyLeuVal	220
Db	635	TGGGCACACTCCAGGTATTTGACCAAGCATGTGAGGACAGCTCTAGGACGGCGCTTGA	694
Qy	221	ProGlyGlySerAlaLeuThrTyrPheSerHisTyrGlnGlnArgLeuAsnSerGlnGln	240
Db	695	CCGGGTGGCACTGCCCTCCACTGGCGCAGCCATACCAGAGAGACTGAATCTCCAGACAG	754
Qy	241	SerCysLeuAsnGlnTyrPhePheAlaMetAlaAspLeuGlnSerLeuAsnProProSerAla	260
Db	755	AGCTCCCTCAATGTAGTGAAGCGCTATGGCGGACTGTGAGTCTGTGGGGCTCCACAGCGC	814
Qy	261	GlnProGlyGlySerSerGlnGlnGlnGlnMetGlnGlnAlaIleArgAlaGlnLeuThr	280
Db	815	GAGCCTGGCGGGTCTCTCAAGAACAGACATATGGACAGCGCATCCGTCTAGCTGATGG	874
Qy	281	LysValLeuAspValSerAspLeuGlnSerValThrSerLysGlnLLeaArgGlnAlaLeu	300
Db	875	AAAGTGTGATGTCACTGACTGTGAAGAGTGTCACTTCCAAAGAGATCCGCCAGGCTCTG	934
Qy	301	GlnLeuArgLeuGlyLeuProLeuGlnGlnTyrArgAspPheLeaAspAsnIleMetLeu	320
Db	935	GAGCTGCCCTGGGGCTCCCTCCACACATCCGTCGACTTATCGAACACAGATGGTGG	994
Qy	321	LeuLeuValAlaGlnArgAspArgAlaSerArgLLeuPheProHisLeuTyrLeuLysSer	340
Db	995	CTGCTGGTGGACACAGCGGAGCCAGGCTCCCGCATCTTCCCCACTCTACTCTGGGCTCA	1054
Qy	341	GlnTyrPheAlaAlaAsnLeuGlnGlnGlnGlnArgAsnArgValThrHisIleLeuAsn	360
Db	1055	GAGTGAACCGCAGCAAACTGGAAGAGCTCAGAGGACAGGGGTACCACCATCTTTGAAC	1114
Qy	361	MetAlaArgGlnLLeaAspAsnPheTyrProGlnArgPheThrTyrHisAsnValArgLeu	380
Db	1115	ATGGCCCGGAGATTGTGAACATTTCTACCTCCAGCGGCTTACCTACACAAATGTGGCGCTC	1174
Qy	381	TrpAspGlnLysSerArgLLeuLeuProHisTyrLysGlnThrHisArgPheLeuGln	400
Db	1175	TGGGATGGAGGTGGCCCAAGCTGTCTGCCACTGGAAGAGAGCAGCAGCGCTTCAATTAG	1234
Qy	401	AlaAlaArgAlaGlnGlyThrHisValLeuValHisCysLysMetGlyValSerArgSer	420
Db	1235	GCTGCAAGACACAGAGGACCCACCAAGTCTGTGTCACTGAAGATGGCGCTCAGCGGCTCA	1294
Qy	421	AlaAlaThrValLeuAlaTyrAlaMetLysGlnTyrGlnCysSerLeuGlnAlaLeu	440
Db	1295	GGGGCCACAGTGTGGCTATGGCATGAACAGTACGAATGACAGCTGTGAGCAGCGCTTG	1354
Qy	441	ArgHisValGlnGlnLeuArgProLLeaAlaArgProAsnProGlyPheLeuArgGlnLeu	460

Db 1355 CGCCACGTCGACGAGCTCCGGCCCATCCGCCGCCCAACCTCGCTTCGCGCAGCTG 1414
 Oy 461 GlnIleTyrgInglYleuThraAlaArgThr 471
 Db 1415 CAGATCTACAGGCGATCTGACGCGCAGAAC 1447
 RESULT 2
 ABO73249
 ID ABO73249 standard; cDNA: 2704 BP.
 XX ABO73249;
 AC ABO73249;
 XX 30-SEP-2002 (first entry)
 DT
 DE Human MAP kinase phosphatase splice form 1 cDNA sequence SEQ ID NO:1.
 XX
 KW Human; phosphatase; mitogen activated protein kinase phosphatase;
 KM MAP kinase; enzyme; chromosome 11; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT 5'UTR 1..93
 FT CDS /*tag= a
 FT 94..1509
 FT /*tag= b
 FT /product= "MAP kinase phosphatase splice form 1"
 FT 1510..2704
 FT 3'UTR /*tag= c
 FT
 TT
 TN
 PN MO200242436-A2.
 XX
 PD 30-MAY-2002.
 XX
 PF 07-NOV-2001; 2001WO-US42995.
 XX
 PR 20-NOV-2000; 2000US-0715177.
 PR 18-JAN-2001; 2001US-0761640.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Wei M, Ketchum KA, Di Francesco V, Beasley EM;
 XX
 DR WPI: 2002-575237/61.
 DR P-PSDB: ABP51653.
 XX
 PT Novel isolated human phosphatase peptide useful for treating disorder
 PT characterized by absence of, inappropriate or unwanted expression of
 PT the phosphatase protein, and as immunogens to raise antibodies -
 XX
 PS Claim 1; Fig 1A; 85pp; English.
 XX
 CC The present invention describes an isolated human phosphatase peptide
 CC (I). (I) can be used for identifying a modulator of (I) by contacting
 CC (I) with an agent and determining if the agent has modulated the
 CC function or activity of (I). (I) is useful for identifying an agent that
 CC binds to (I). (I) with an agent and assaying the contacted
 CC mixture to determine whether a complex is formed with the agent bound
 CC (I). The human phosphatases from the present invention are mitogen
 CC activated protein (MAP) kinase phosphatases. These human MAP kinase
 CC phosphatases are located on chromosome 11. (I) and the polynucleotide
 CC sequences encoding (I) can be used in gene therapy. The present sequence
 CC encodes human MAP kinase phosphatase splice form 1 from the present
 CC invention.
 CC
 SQ Sequence 2704 BP: 569 A; 874 C; 794 G; 467 T; 0 other:
 Alignment Scores:
 Pred. No.: 4.71e-195 Length: 2704
 Score: 2436.00 Matches: 471
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0

DB: 24 Gaps: 0
 US-09-761-640-4 (1-471) x ABO73249 (1-2704)
 Oy 1 MetalAlaLeuValThrValSerArgSerProProGlySerGlyAlaSerThrProValGly 20
 Db 94 ATGGCCCTGTCTACAGTAGAGCGGTTCGCCCCCGGCGACGGCGGCTCCACGCCCTGGGG 153
 Oy 21 ProTrpAspGlnAlaValGlnArgArgSerArgLeuGlnArgGlnSerPheAlaVal 40
 Db 154 CCTGGGACCAAGGCGGTCCAGCAGAGAGTGCACCTCCAGCAAGGCAAGCTTTCGGGTG 213
 Oy 41 LeuArgGlyAlaValLeuGlyLeuGlnAspGlyGlyAspAspAspAlaAlaGlyAla 60
 Db 214 CTCCTGGGGCTGTCTCTGGACTGAGATGGAGGGGCAATGATGATGCACAGAGGCC 273
 Oy 61 SerSerGluProThrGluAlaProSerGluGluGluGluGluGluGluGluGluGluGlu 80
 Db 274 AGTCTGAGCCACACAGAGAGAGGCCCGAGTGAGAGAGCTCCACGGGGACAGACAC 333
 Oy 81 PheGlyGlnGlySerGlnSerProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 100
 Db 334 TTCGGCAAGAGATCCAGAGTCCAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 393
 Oy 101 MetValGlnLeuLeuArgProGlnAspAspIleArgLeuAlaGlnLeuGluAlaPro 120
 Db 394 ATGTGTACAGCTGTCTGAGCCCGCAGATGACATCCGCTGGCAGCCACTGGAGGCAACC 453
 Oy 121 ArgProProArgLeuArgTyrLeuLeuValValSerThrArgGluGluGluGluGluSer 140
 Db 454 CGGCTCCCGCGGCTCGCTACCTGCTAGTTCTTACACAGAGAGAGAGAGAGAGAGAGAG 513
 Oy 141 GlnAspGluThrValLeuLeuGlyAlaAspPheProAspSerSerSerProSerCysThr 160
 Db 514 CAGGATGAGAGGCTCTCTGGCGTGGATTTCTTGACAGAGCTCCCGACGTCGACCC 573
 Oy 161 LeuGlyLeuValLeuProLeuTrpSerAspThrGlnValTyrLeuAspGlyAspGly 180
 Db 574 CTGGGCTGTGTTCTTCCCTCTGAGTGCACACCCAGGTGTACTTGTGATGGAGCGGGGCG 633
 Oy 181 PheSerValThrSerGlyGlyGlnSerArgIlePheLysProIleSerIleGlnThrMet 200
 Db 634 TTCAGCTGAGCTGTGGTGGGCAAGCCGATCTTCAAGCCCATCTCATCCAGACATG 693
 Oy 201 TrpAlaThrLeuGlnValLeuHisGlnAlaCysGluAlaAlaLeuGlySerGlyLeuVal 220
 Db 694 TGGGCCACACTCCAGGATTTGACCAACAGCATGTGAGGAGCTTAAGGAGGCGCTTGTGA 753
 Oy 221 ProGlyGlySerAlaLeuThrTrpAlaSerHisTyrGlnGluArgLeuAsnSerGluGln 240
 Db 754 CCGGCTGGCAGTGGCTCTCACTGGGCGCAGCCACTACAGAGAGAGCTGAACCTCCGAACG 813
 Oy 241 SerCysLeuAsnGluTyrTrpThrAlaMetAlaAspLeuGluSerLeuArgProProSerAla 260
 Db 814 AGCTCCCTCAATGAGTAGAGCGCTATGGCCGACCTGAGATCTCTCGGCTCCAGCGCC 873
 Oy 261 GluProGlyGlySerSerGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 280
 Db 874 GAGCTTGGCGGGTCTCTCAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 933
 Oy 281 LysValLeuAspValSerAspLeuGluSerValThrSerLysGluIleArgGlnAlaLeu 300
 Db 934 AAGCTGTGGATGTACAGTACCTGGAGAGTGTCACTTCCAAAGAGATCCGCAAGGCTGTG 993
 Oy 301 GluLeuArgLeuGlyLeuProLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 320
 Db 994 GAGCTGGCGCTGGGGCTCCCTCCAGCAGTACCTGACTCATGAGCAACAGATGCTG 1053
 Oy 321 LeuLeuValAlaGlnArgAspArgAlaSerArgIlePheProHisLeuTyrLeuGlySer 340
 Db 1054 CTGCTGTGTGGACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1113
 Oy 341 GluTrpAsnAlaAlaAsnLeuGluGluLeuGlnArgAsnArgValThrHisIleLeuAsn 360

```
Db 1114 GAGTGAACGACGCAACCTGAGAGCTGACGAGAAAGGCTACCATTTTGAAC 1173
Qy 361 MetAlaArgGluLeuAspAspPheTyrProGluArgPheThrTyrHisAsnValArgLeu 380
Db 1174 ATGGCCGGGAGATTGACACTTCTACCTGAGCGCTTCACTACCAATGTCGGCTC 1233
Qy 381 TrpAspGluSerAlaGlnLeuLeuProHisTyrPlySerGluThrHisAspPheIleGlu 400
Db 1234 TGGGATGAGAGTGGCCGACGCTGCTCCGACACGAGAGAGACGCCACCTTCATTGAG 1293
Qy 401 AlaAlaArgAlaGlnGlyThrHisValLeuValHisCysLysMetGlyValSerArgSer 420
Db 1294 GCTGCAAGACACAGGGGACCCAGCTGCTGTCACAGTGAAGTGGCGTCAGCCGCTCA 1353
Qy 421 AlaAlaThrValLeuAlaTyrAlaMetLysGlnTyrGlnGlySerLeuGlnAlaLeu 440
Db 1354 GGGGCGACACTGCTGGCTTGTCCATGACCACTACCAATGACAGCTGGAGAGGCCCTG 1413
Qy 441 ArgHisValGlnGluLeuArgProIleAlaArgProAsnProGlyPheLeuArgGlnLeu 460
Db 1414 CGCCAGTGCAGAGCTCCGCGCCATGCGCCGACCAACCTGGCTTCTGGCCAGCTG 1473
Qy 461 GlnIleTyrGlnGlyIleLeuThrAlaArgThr 471
Db 1474 CAGATCTACAGGGCATCTCTGACGGCAGAAC 1506

RESULT 3
AB073252
ID AB073252 standard; cdna; 2704 BP.
XX
AC AB073252;
XX
DT 30-SEP-2002 (first entry)
XX

Human MAP kinase phosphatase splice form 1 cDNA sequence SEQ ID NO:7.
DE Human MAP kinase phosphatase splice form 1 cDNA sequence SEQ ID NO:7.
XX Human; phosphatase; mitogen activated protein kinase phosphatase;
KW MAP kinase; enzyme; chromosome 11; single nucleotide polymorphism;
KW SNP; gene; ss.
XX
OS Homo sapiens.
XX
FH Key 1. 93 Location/Qualifiers
FH 5'UTR /tag= a
FH CDS 94..1509
FH /tag= b
FH /product= "MAP kinase phosphatase splice form 1"
FH variation replace(577,A)
FH /tag= c
FH variation /standard_name= "single nucleotide polymorphism (SNP)"
FH /tag= d
FH 3'UTR /standard_name= "single nucleotide polymorphism (SNP)"
FH /tag= e
FH variation replace(2641,A)
FH /tag= f
FH /standard_name= "single nucleotide polymorphism (SNP)"
XX
FN W0200242436-A2.
XX
PD 30-MAY-2002.
XX
PF 07-NOV-2001; 2001WO-US42995.
XX
PR 20-NOV-2000; 2000US-0715177.
XX
PR 18-JAN-2001; 2001US-0761640.
XX (PEKE ) PE CORP NY.
XX
PA Wei M, Ketchum KA, Di Francesco V, Beasley EM.
XX
PI
```

```
XX
DR WPI: 2002-575237/61.
DR P-PSDB; ABP51653.
XX
PT Novel isolated human phosphatase peptide useful for treating disorder
PT characterized by absence of, inappropriate or unwanted expression of
PT the phosphatase protein, and as immunogens to raise antibodies
XX
PS Claim 1; Fig 3A; 85pp; English.
XX

The present invention describes an isolated human phosphatase peptide
CC (I) (I) can be used for identifying a modulator of (I) by contacting
CC (I) with an agent and determining if the agent has modulated the
CC function or activity of (I). (I) is useful for identifying an agent that
CC binds to (I), by contacting (I) with an agent and assaying the contacted
CC mixture to determine whether a complex is formed with the agent bound
CC (I). The human phosphatases from the present invention are mitogen
CC activated protein (MAP) kinase phosphatases. These human MAP kinase
CC phosphatases are located on chromosome 11. (I) and the polynucleotide
CC sequences encoding (I) can be used in gene therapy. The present sequence
CC encodes human MAP kinase phosphatase splice form 1 from the present
CC invention.
XX
SQ Sequence 2704 BP: 569 A; 874 C; 794 G; 467 T; 0 other:
XX

Alignment Scores:
Pred. No.: 4.71e-195 Length: 2704
Score: 2436.00 Matches: 471
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 24 Gaps: 0

US-09-761-640-4 (1-471) x AB073252 (1-2704)
Qy 1 MetalAlaLeuValThrValSerArgSerProProGlySerGlyAlaSerThrProValGly 20
Db 94 ATGGCCCTGTGCACAGTACAGCGTTCGCCGCCGCGGACGCGGCTCCACCCGCTGGGG 153
Qy 21 ProTrpAspGlnAlaValGlnArgArgSerArgLeuGlnArgArgGlnSerPheAlaVal 40
Db 154 CCCTGGGACAGGCGGCTCCAGCGAAGAGTGCAGCTCCAGCAAGAGGAGAGCTTGGCGGTG 213
Qy 41 LeuArgGlyAlaValLeuGlyLeuGlnAspGlyGlyAspAsnAspAlaAlaGlnAla 60
Db 214 CTCGGTGGGCTGTCTCTGGAGCTGACAGATGAGAGGAGCAATGATGACGACGAGGCC 273
Qy 61 SerSerGluProThrGlnLysAlaProSerGlnGlnLeuHisGlyAspGlnThrAsp 80
Db 274 AGTCTGAGCCACAGAGAGGCCCGAGTGAAGAGGCTCCAGGGGAGACAGACGAC 333
Qy 81 PheGlyGlnGlySerGlnSerProGlnLysGlnGlnLeuGlnArgGlnHisLeuHisLeu 100
Db 334 TTGGGGCAGAGATCCAGAGTCCCAAGACAGAGAGAGAGGAGGACACTGCACTC 393
Qy 101 MetValGlnLeuLeuArgProGlnAspAspIleArgPheAlaGlnLeuGlnAlaPro 120
Db 394 ATGCTACAGCTGCTGAGGCGGAGAGATCAGCTGCGACAGCCAGCTGGAGGACCC 453
Qy 121 ArgProProArgLeuArgTyrLeuLeuValAlaSerThrArgGlnGlyGlnLeuSer 140
Db 454 CGGCTTCGCCGCTCCCTACCTGCTGTAGTTTACAGAGAGAGAGAGAGAGTCTGAGG 513
Qy 141 GlnAspGluThrValLeuLeuGlnGlyValAspPheProAspSerSerProSerCysThr 160
Db 514 CAGAGTACAGAGGCTCTCTCGGGCGTGAGATTCTCTGACAGCAGCTCCACCTGACAC 573
Qy 161 LeuGlyLeuValLeuProLeuThrSerAspThrGlnValTyrLeuAspGlyGly 180
Db 574 CTGGGCTGTGTGGCCCTCTGAGTGAACACCAGAGTGTACTTAAATGACAGGGGGCC 633
Qy 181 PheSerValThrSerGlyGlnSerArgGlyPheLysProIleSerIleGlnThrMet 200
```

```

Db 634 TTCAGCGTCAGCTCTGTGGGCAAAACCGGATTCACAGCCCATCTCCATCCAGACATG 693
OY 201 TTPAlaThrLeuGlnValLeuHisGlnAlaCysGlnAlaLeuGlnSerGlyLeuVal 220
Db 694 TGGGCCACACTCCAGGATTTGCAACAGCATGTGAGGCGAGCTCTAGCGAGCGCTGTGA 753
OY 221 ProGlyGlySerAlaLeuThrTTPAlaSerHisTyrGlnGlnIleArgLeuAsnSerGlnGln 240
Db 754 CCGGGGTGGCAGTGGCTTCACCTAGGCGCACCATACAGAGAGAGACTGAAGTCCGACAG 813
OY 241 SerCysLeuAsnGlnTTPThrAlaMetAlaAspLeuGlnSerLeuArgProProSerAla 260
Db 814 AACTGCTCTCAATGAGTGGAGGCTATGGCCGACCTGGAGTCTCTGGGCTCCGACGGCC 873
OY 261 GluProGlyGlySerSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 280
Db 874 GAGCTGGGGGCTCTTCACACAGAGGAGATGAGACAGGCGATCCGTCTGAGCTGTGG 933
OY 281 LysValLeuAspValSerAspLeuGlnSerValThrSerGlyLeuArgGlnAlaLeu 300
Db 934 AAAGTGTGATGTACGTACCTGGAGAGTGTACCTCCAAAGAGATCCGCGAGGCTGTG 993
OY 301 GluLeuArgLeuGlyLeuProLeuGlnGlnTyrArgAspPheIleAspAsnGlnMetLeu 320
Db 994 GAGCTGGCGCTGGGCTCCCTCCAGCAGTACCGTGAATCTTCAACACAGATGCTG 1053
OY 321 LeuLeuValAlaGlnArgAspArgAlaSerArgIlePheProHisLeuTyrLeuGlySer 340
Db 1054 CTGCTGGTGGCAGAGGGAGCAGGCTCCGATCTCCGACCTCTCACTGAGGCTCTCA 1113
OY 341 GluTPAsnAlaAlaAsnLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 360
Db 1114 GAGTGGAAACGAGCAAACTGGAGGAGTGCAGAGAGCAAGGTTACCCACATCTTGAAAC 1173
OY 361 MetAlaArgGlnIleAspAsnPheTyrProGlnIleArgPheThrTyrHisAsnValArgLeu 380
Db 1174 ATGGCCCGGGAATTGACAACTTACCTCGAGCGCTTACCTACACAAATGTGGCCCTTC 1233
OY 381 TTPAspGlnGlnSerAlaGlnLeuLeuProHisTyrPlyGlnThrHisArgPheIleGln 400
Db 1234 TGGGATGAGAGATCGCGCCAGCTGTGCGCCAGCTGGAAGAGACGACGCGCTTCATTGAG 1293
OY 401 AlaAlaArgAlaGlnGlyThrHisValLeuValHisCysLysMetGlyValSerArgSer 420
Db 1294 GCTGCAGAGACAGAGGCGACCCAGCTGTGTCACACTGCAGATGGCGCTCAGACCCCTCA 1353
OY 421 AlaAlaThrValLeuAlaThrAlaMetLysGlnTyrGlnLysSerLeuGlnAlaLeu 440
Db 1354 GCGGCCACAGTGTGCTGCTATGCCATGAAGCAGTACGAATGACGCTGGAGCGGCCCTG 1413
OY 441 ArgHisValGlnGlnLeuArgProIleAlaArgProAsnProGlyPheLeuArgGlnLeu 460
Db 1414 CCGCAGCTGACAGAGACTCGCGGCCATCGCCGCCCAACCTGTGCTTCTGCGCCAGCTG 1473
OY 461 GlnIleTyrGlnGlyIleLeuThrAlaArgThr 471
Db 1474 CAGATCTACAGGAGGATCTGACGCGCAGACACC 1506

```

RESULT 4
AAD36061
ID AAD36061 standard; cDNA, 2718 BP.

AC AAD36061;

DT 09-AUG-2002 (first entry)

DE Human dual-specificity phosphatase 15 (DSP-15) cDNA.

KW Human: dual-specificity phosphatase 15; DSP-15; antiallergic; cytostatic;
KW immunosuppressive; MAP; mitogen activated protein kinase; cancer; enzyme;
KW signal transduction; cell proliferation; Duchenne muscular dystrophy;
KW cell cycle abnormality; graft-versus-host disease; autoimmune disease;
KW metabolic disease; allergy; screening; chromosome 11q; gene; ss.

```

XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX CDS 35..2014
XX FT /*tag=
XX FT /product="Human DSP-15 protein"
XX PN MO200224740-A2.
XX PD 28-MAR-2002.
XX PF 19-SEP-2001; 2001WO-US29406.
XX PR 19-SEP-2000; 2000US-233833P.
XX PR 18-SEP-2001; 2001US-0955732.
XX PA (CEPT-) CEPTYR INC.
XX PI Luche RM, Wel B;
XX P1 WPL: 2002-394127/42.
XX P2 P-PSDB: AAE22729.
XX PT New dual-specificity phosphatase 15 polypeptide and polynucleotides,
XX PT useful for treating e.g. Duchenne muscular dystrophy, cancer,
XX PT graft-versus-host disease, autoimmune diseases, allergies, metabolic
XX PT diseases
XX PS Claim 7; Fig 1; 91pp; English.
XX CC The invention relates to a new isolated dual-specificity phosphatase 15
XX CC (DSP-15) polypeptide which retains the ability to dephosphorylate an
XX CC activated MAP (mitogen activated protein) kinase. DSPs are phosphatases
XX CC that dephosphorylate both phosphotyrosine and phosphothreonine/serine
XX CC residues. DSP-15 polypeptides may be used to identify agents that
XX CC modulate DSP-15 activity, where such agents may inhibit or enhance signal
XX CC transduction via a MAP-kinase cascade, leading to cell proliferation. DSP
XX CC polypeptides, modulating agents, and/or polynucleotides encoding the
XX CC polypeptides may be used to modulate DSP-15 activity in a patient, and to
XX CC ameliorate disorders such as Duchenne muscular dystrophy, cancer, graft-
XX CC versus-host disease, autoimmune diseases, allergies, metabolic diseases,
XX CC abnormal cell growth, abnormal cell proliferation and cell cycle
XX CC abnormalities. DSP-15 alternate form polypeptides are useful in screening
XX CC assays for modulators of enzyme activity and/or substrate binding. The
XX CC present sequence is human DSP-15 cDNA. Human DSP-15 gene is located on
XX CC chromosome 11q.
XX SQ Sequence 2718 BP; 559 A; 893 C; 795 G; 471 T; 0 other;
XX
Alignment Scores:
Pred. No.: 3 29e-194 Length: 2718
Score: 2426.00 Matches: 469
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.59% Indels: 0
DB: 24 Gaps: 0

```

US-09-761-640-4 (1-471) x AAD36061 (1-2718)

```

OY 1 MetAlaLeuValThrValSerArgSerProGlySerGlyAlaSerThrProValGly 20
Db 35 ATGGCCCTGCTGTCACAGTGAAGCGTTCGCCCGGCGAGCGGCGCTTCACAGCCCGTGGGG 94
OY 21 ProTPAspGlnAlaValAlaGlnArgArgSerArgLeuGlnIleArgGlnSerPheAlaVal 40
Db 95 CCTGGGACCGAGCGGTTCACGACGAGAGTCACTCCAGCAGAGGCGAGGCTTGGGGTG 154
OY 41 LeuArgGlyAlaValLeuGlyLeuGlnAspGlyGlyAspAsnAspAspAlaAlaGlnAla 60
Db 155 CTCCTGGGGGCTCTCTGGGACTGCAAGATGAGGGGACAAATATATATGACAGAGGCC 214
OY 61 SerSerGluProThrGlnLysAlaProSerGlnGlnGlnLeuHisGlnLysGlnThrAsp 80

```

[illegible]

Db	1295	GCGGCCACAGTGTGGCCCTAAGCCATGACAGCAATGACAGCTTGAGCAGGCCCCCTG	1354
Qy	441	ArghisValGInGInGInLeuArgProIleAlaArgProAsnProGlyIleuArgGInleu	460
Db	1355	CGCCACGTGACGAGAGCTCCGGGCCCATGCGCCGCCCAACCTTGCTCTCGCCAGCTTG	1414
Qy	461	GInIleTyrgInGInGlyIleLeuThrAla	469
Db	1415	CAGATCTACACGAGGCATCTGACGCGC	1441
RESULT 5			
AAF30485	ID	AAF30485 standard; cDNA; 2781 bp.	
XX	AC	AAF30485;	
XX	DT	29-MAY-2001 (first entry)	
XX	DE	Human protein phosphatase and kinase protein-10 cDNA 5039718C81.	
XX	KW	Protein phosphatase and kinase protein; PPKP-10; human;	
KW	KW	gastrointestinal disorder; immune system disorder;	
KW	KW	neurological disorder; cell proliferative disorder; cancer;	
XX	OS	diagnosis; therapy; ss.	
XX	XX	Homo sapiens.	
FH	FH	Key	Location/Qualifiers
FT	FT	CDS	87..2066
FT	FT		/*tag= a
FT	FT	sig_peptide	87..143
FT	FT		/*tag= b
FT	FT	mat_peptide	144..2063
FT	FT		/*tag= c
FT	FT	misc_feature	219..279
FT	FT		/*tag= d
FT	FT	misc_feature	921..980
FT	FT		/note= "unique fragment"
XX	XX		/*tag= e
XX	XX		/note= "unique fragment"
PN	PN	WO200120004-A2.	
PD	PD	22-MAR-2001.	
XX	XX	14-SEP-2000; 2000WO-US25515.	
PF	PF		
XX	XX	15-SEP-1999; 99US-0154141.	
PR	PR		
XX	XX	(INCYTE GENOMICS INC.	
PA	PA		
XX	XX	Yue H, Tang YT, Bandman O, Hillman JL, Baughn MR, Azimzal Y;	
PI	PI	Lu DAM;	
PI	PI		
XX	XX	WPI: 2001-244811/25.	
DR	DR	P-PSDB: AAB20331.	
XX	XX		
PT	PT	Novel human protein phosphatase and kinase proteins for diagnosis,	
PT	PT	treatment and prevention of gastrointestinal, immune system,	
PT	PT	neurological and cell proliferative disorders -	
XX	XX		
PS	PS	Claim 5; Page 101-102; 103pp; English.	
XX	XX		
CC	CC	The present sequence is that of cDNA encoding novel human	
CC	CC	protein phosphatase and kinase protein PPKP-10 (see AAB20331).	
CC	CC	The cDNA was initially identified in Incyte Clone ID No. 5039718C81,	
CC	CC	from a colon tumour tissue library. Tissues that express PPKP-10	
CC	CC	(as a fraction of total tissues expressing PPKP-10) include	
CC	CC	reproductive (0.343), gastrointestinal (0.194) and haematopoietic	
CC	CC	or immune (0.134). Diseases or conditions associated with tissues	
CC	CC	expressing PPKP-10 (as a fraction of total tissues expressing	
CC	CC	PPKP-10) include cancer (0.552), inflammation or trauma (0.314) or	
CC	CC	cell proliferation (0.090). The encoded protein shows homology to	

CC Drosophila melanogaster MAP kinase phosphatase. The invention
 CC provides human PPKP-1 to -11 polypeptides (see AAB20322-32) and
 CC polynucleotides (see AAF30476-86). It also provides expression
 CC vectors, host cells, antibodies, agonists and antagonists, as well
 CC as methods for diagnosing, treating or preventing disorders
 CC associated with expression of PPKP, including gastrointestinal
 CC disorders, immune system disorders, neurological disorders and cell
 CC proliferative disorders, including cancer.

XX Sequence 2781 BP: 576 A: 906 C: 820 G: 479 T: 0 other:

Alignment Scores:

Pred. No.:	3,39e-194	Length:	2781
Score:	2426.00	Matches:	469
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	99.59%	Indels:	0
	22	Gaps:	0

US-09-761-640-4 (1-471) x AAF30485 (1-2781)

Oy 1 MetAlaLeuValThrValSerArgSerProGlySerGlyAlaSerThrProValGly 20
 Db 87 ATGGCCCTGGTACAGTAGCCGTTCCGCCCGGCGAGCGGCGCTCCAGCGCCGTGGGG 146
 Oy 21 ProTPASPGLuAlaValGlnArgArgSerArgLeuGlnArgArgGlnSerPheAlaVal 40
 Db 147 CCTGGGACACAGCGGCTCCAGGAGAGAGTGCACCTCCAGGAGGACAGCTTTGGCGTG 206
 Oy 41 LeuArgGlyAlaValLeuGlyLeuGlnAspGlyGlyAspAsnAspAspAlaAlaGluAla 60
 Db 207 CTCCGTGGGGGCTGCTCGGAGCTGAGATGAGAGGACAAATGATGATGCACAGAGGCC 266
 Oy 61 SerSerGluProThrGlnGlyAlaProSerGluGluGluLeuHisGlyAspGlnThrAsp 80
 Db 267 AGTTCTGAGCCCAACAGAAAGGCCCGGAGTGAGAGAGCTCCAGGGGACACAGACAGAC 326
 Oy 81 PheGlyGlnGlySerGlnSerProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 100
 Db 327 TTCCGGGACAGAGATCCACAGAGTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 386
 Oy 101 MetValGlnLeuLeuArgProGlnAspAspIleArgLeuAlaAlaGlnLeuGluAlaPro 120
 Db 387 ATGTACAGAGCTCTGAGGCGGAGATGATGATGATGATGATGATGATGATGATGATGAT 446
 Oy 121 ArgProProArgLeuArgTyrLeuLeuValValSerThrArgGluGlyGluGlyLeuSer 140
 Db 447 CGGCTCCCGGCTCCGCTACCTGCTGATGTTCTTACACGAGAGAGAGAGAGAGAGAG 506
 Oy 141 GlnAspGluThrValLeuLeuGlyValAspPheProAspSerSerSerProSerCysThr 160
 Db 507 CAGGATGAGAGAGGCTCTCTGGGCGCTGATTTCCCTGACAGAGAGAGAGAGAGAGAG 566
 Oy 161 LeuGlyLeuValLeuProLeuTrpSerAspThrGlnValTyrLeuAspGlyAspGlyGly 180
 Db 567 CTGGGCTGCTGCTTCCCTCGAGTGACACCAAGGTGATGATGATGATGATGATGATGAT 626
 Oy 181 PheSerValThrSerGlyGlyGlnSerArgIlePheLeysProIleSerIleGlnThrMet 200
 Db 627 TTCACAGCTGACCTGCTGGGCAAGCGGATCTCAAGCCATCTCCATCCAGACAGAG 686
 Oy 201 TrpAlaThrLeuGlnValLeuHisGlnAlaCysGlnAlaAlaLeuGlySerGlyLeuVal 220
 Db 687 TGGGCGACACTCCAGATATGGACCAAGCATGTGAGGCGCTTGTAGGCGGCGCTGTGA 746
 Oy 221 ProGlyGlySerAlaLeuThrTrpAlaSerHisTyrGlnGluArgLeuAsnSerGluGln 240
 Db 747 CCGGGGAGAGTCCCTACCTGAGGCGAGCCACTACAGAGAGAGAGAGAGAGAGAGAGAG 806
 Oy 241 SerCysLeuAsnGluTrpThrAlaMetAlaAspLeuGluSerLeuArgProProSerAla 260
 Db 807 AGCTGCTCAATGAGTGGAGGCTATGGCCGACGTGAGAGTCTCTGCGGCTCCAGTGGCC 866

Oy 261 GluProGlyGlySerSerGluGlnGlnGlnMetGluGlnAlaIleArgAlaGluLeuTrp 280
 Db 867 GAGCTCGGGGCTCTCTCAAAAGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 926
 Oy 281 LysValLeuAspValSerAspLeuGluSerValThrSerLysGluIleArgGlnAlaLeu 300
 Db 927 AAGGTGTGATGTCAGTACCTGAGAGAGTGCACCTTCCAAAGAGATCCGAGGCTGTG 986
 Oy 301 GluLeuArgLeuGlyLeuProLeuGlnGlnTyrArgAspPheIleAspAsnIleLeu 320
 Db 987 GAGCTGGCGCTGGGCTCCCTCCAGAGTACCGGATGATGATGATGATGATGATGATGAT 1046
 Oy 321 LeuLeuValAlaGlnArgAspArgAlaSerArgIlePheProHisLeuTyrLeuGlySer 340
 Db 1047 CTGCTGGTGGCACAGCGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1106
 Oy 341 GluTrpAsnAlaAlaAsnLeuGlnGluLeuGlnGlnArgAsnArgValThrHisIleLeuAsn 360
 Db 1107 GAGTGGAAAGCGACCAACCTGAGAGAGCTGACAGAGAGAGAGAGAGAGAGAGAGAGAG 1166
 Oy 361 MetAlaArgGluIleAspAsnPheTyrProGluArgPheThrTyrHisAsnValArgLeu 380
 Db 1167 ATGGCCCGGAGATTTGACAACTTACCTGAGCCCTTACCTACACACATGTGGCGCTTC 1226
 Oy 381 TrpAspGluGluSerAlaGlnLeuLeuProHisTyrPlysGluThrHisArgPheIleGlu 400
 Db 1227 TGGGATGAGAGATCGGCCAGCTGCTGCCGACCTGGAAGAGAGAGAGAGAGAGAGAG 1286
 Oy 401 AlaAlaArgAlaGlnGlyThrHisValLeuValHisCysLysMetGlyValSerArgSer 420
 Db 1287 GCTGCAAGAGCACAGAGGACCCAGCGTGTGCTCACTGCAAGATGGGCGTACGCGCTCA 1346
 Oy 421 AlaAlaThrValLeuAlaTyrAlaMetLysGlnTyrGluCysSerLeuGlnAlaLeu 440
 Db 1347 GCGGCACAGTCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1406
 Oy 441 ArgHisValGlnGluLeuArgProIleAlaArgProAsnProGlyPheLeuArgGlnLeu 460
 Db 1407 CGCCACGTGACAGAGAGCTCGGCCATCGGCCGCCCAACACCTGGCTTCTGGCCAGCTG 1466
 Oy 461 GlnIleTyrGlnGlyIleLeuThrAla 469
 Db 1467 CAGATCTACAGAGGATCTCTGACGCC 1493

RESULT 6
 ABQ73250
 ID ABQ73250 standard: cDNA: 2852 BP.
 AC ABQ73250;
 XX 30-SEP-2002 (first entry)
 DT
 DE Human MAP kinase phosphatase splice form 2 cDNA sequence SEQ ID NO:2.
 DE Human: phosphatase; mitogen activated protein kinase phosphatase;
 KW MAP kinase; enzyme; chromosome 11; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FH 1..56
 FT 5'UTR
 FT
 FT CDS
 FT 57..2036
 FT /tag= a
 FT /tag= b
 FT /product= "MAP kinase phosphatase splice form 2"
 FT 2037..2852
 FT /tag= c
 XX
 PN WO200242436-A2.
 XX
 XX 30-MAY-2002.
 PD
 XX 07-NOV-2001; 2001WO-USA2995.
 PF

XX 20-NOV-2000; 2000US-0715177.
 PR 18-JAN-2001; 2001US-0761640.
 XX (PEKE) PE CORP NY.
 PA
 PI Wei M, Ketchum KA, Di Francesco V, Beasley EM;
 XX
 DR WPI: 2002-575237/61.
 XX P-PSDB; ABP51654.
 XX
 PT Novel isolated human phosphatase peptide useful for treating disorder
 PT characterized by absence of, inappropriate or unwanted expression of
 PT the phosphatase protein, and as immunogens to raise antibodies
 XX
 PS Claim 1, Fig 1B; 85pp; English.

CC The present invention describes an isolated human phosphatase peptide
 CC (I). (I) can be used for identifying a modulator of (I) by contacting
 CC (I) with an agent and determining if the agent has modulated the
 CC function or activity of (I). (I) is useful for identifying an agent that
 CC binds to (I), by contacting (I) with an agent and assaying the contacted
 CC mixture to determine whether a complex is formed with the agent bound
 CC (I). The human phosphatases from the present invention are mitogen
 CC activated protein (MAP) kinase phosphatases. These human MAP kinase
 CC phosphatases are located on chromosome 11. (I) and the polynucleotide
 CC sequences encoding (I) can be used in gene therapy. The present sequence
 CC encodes human MAP kinase phosphatase splice form 2 from the present
 CC invention.

XX SQ Sequence 2852 BP: 674 A: 895 C: 807 G: 476 T: 0 other:

Alignment Scores:

Pred. No.:	3 51e-194	Length:	2852
Score:	2426.00	Matches:	469
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	99.59%	Indels:	0
DB:	24	Gaps:	0

US-09-761-640-4 (1-471) x ABQ73250 (1-2852)

QY 1 MetAlaLeuValThrValSerArgSerProGlySerGlyAlaSerThrProValGly 20
 DB 57 ATGGCCCTGGTACAGTAGAGCGGTTCGCCCGGCGAGCGGCGCTCCAGCGCGGGGG 116
 QY 21 ProTPASpGlnAlaValGlnArgSerArgLeuGlnArgGlnSerPheAlaVal 40
 DB 117 CCCTGGACACAGCGGCTCCAGGAGAGTCCAGTCCAGCCGAAGGCAAGCTTGGCGTG 176
 QY 41 LeuArgGlyAlaValLeuGlyLeuGlnAspGlyGlyAspAsnAspAspAlaGlnAla 60
 DB 177 CTCCTGGGGGCTGCTCTGGAGTGGAGTGGAGGAGCAATGATGATGCACCAAGGCC 236
 QY 61 SerSerGluProThrGlnLysAlaProSerGluGluGlnLysGlnLysAlaSerPhe 80
 DB 237 AGTCTGAGCCAAAGAGAGAGGCCCGAGTGAAGAGAGTCCAGGGGAGCCAGACAGC 296
 QY 81 PheGlyGlnGlySerGlnSerProGlnLysGlnGluGlnArgGlnLysLeuLysLeu 100
 DB 297 TTCGGGCAAGATCCCAAGTCCCAAGAGCAGAGCAGAGCAGACACTGCACCTC 356
 QY 101 MetValGlnLeuLeuArgProGlnAspSerPheLeuAlaGlnLeuGlnAlaPro 120
 DB 357 ATGGTACAGCTGCTGAGGCGCAGAGATGACATCCCTGGGAGCCCACTGGAGGCCACC 416
 QY 121 ArgProPheArgLeuArgTyrLeuLeuValValSerThrArgGluGluGlyLeuSer 140
 DB 417 CGGCTCCCGCGGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 476
 QY 141 GlnAspGluThrValLeuLeuGlyValAspPheProAspSerSerProSerCysThr 160
 DB 477 CAGGATGAGAGCGTCTCTGCGGCTGATTTCCCTGCACAGACTCCCGACGTCGACC 536

QY 161 LeuGlyLeuValLeuProLeuTyrSerAspThrGlnValTyrLeuAspGlyAspGly 180
 DB 537 CTGGCGCTGGTTCGCCCTCTGGAGTGCACCCAGGATGACTTAAAGTGGAGCGGGGG 596
 QY 181 PheSerValThrSerGlyGlyGlnSerArgLlePheLysProLleSerLleGlnThrMet 200
 DB 597 TTCAGCTGACGCTCTGTGGGCAAGCGGATCTTCAAGCCATTCATCCAGACCATG 656
 QY 201 TrpAlaThrLeuGlnValLeuHisGlnAlaCysGluAlaAlaLeuGlySerGlyLeuVal 220
 DB 657 TGGGCCACACTCCAGGATATGCACCAAGCATGTAGAGGACCTTAGGACGGGCTTGA 716
 QY 221 ProGlyGlySerLleAlaThrTrpAlaSerHisTyrGlnGluArgLeuAsnSerGluGln 240
 DB 717 CCGGGTGGCAGTCCCTCAGCTGGGCCAGCAGCTACAGAGAGACTGAATCCGAAACAG 776
 QY 241 SerCysLeuAsnGluThrTrpAlaMetAlaAspLeuGlySerLeuArgProProSerAla 260
 DB 777 AGCTGCCTCAATGAGTGAAGCGGTATGGCCGACCTGAGTCTGCGGCTCCAGCGCC 836
 QY 261 GluProGlyGlySerSerGluGlnGlnMetGluGlnAlaLleArgAlaGluLeuTrp 280
 DB 837 GAGCCTGGCGGCTCTCAGAACAGAGCAGATGAGACAGCGATCCGCTGAGCTGTGG 896
 QY 281 LysValLeuAspValSerAspLeuGlySerValThrSerLysGluLleArgGlnAlaLeu 300
 DB 897 AAGGTTGGATGTCAAGTACCTGGAGAGTGTCACTTCAAAAGATCCGCGCAGCTGTG 956
 QY 301 GluLeuArgLeuGlyLeuProLeuGlnGlnTyrArgAspPheLleAspAsnGlnMetLeu 320
 DB 957 GAGCTCGGCTGGGGGCTCCCGCCAGCAGTACCGTGAATTCATGACACACAGATGCTG 1016
 QY 321 LeuLeuValAlaGlnArgAspArgAlaSerAlaGlyLlePheProHisLeuTyrLeuGlySer 340
 DB 1017 CTGCTGGTGGCACAGGGGAGCGAGCCTCCCGCATCTTCCCGACCTGACCTGGGCTCA 1076
 QY 341 GluTrpAsnAlaAlaAsnLeuGluGlnLeuGlnArgAsnArgValThrHisLleLeuAsn 360
 DB 1077 GAGTGAAGCGACCAAACTGGAGAGCTGAGAGAGGAGAACAGGGTATCCACATCTTGACC 1136
 QY 361 MetAlaArgGluLleAspAsnPheTyrProGluArgPheThrTyrHisAsnValArgLeu 380
 DB 1137 ATGGCCCGGAGATTGACAACTTTCACCTGAGCCCTTACCTACCAATGTGCGGCTC 1196
 QY 381 TrpAspGluGluSerAlaGlnLeuLeuProHisTrpLysGluThrHisArgPheLleGlu 400
 DB 1197 TGGATGAGAGTCCGGCCAGCTGCTGCCGCACTGGAAGAGACGACCGCTCAATTGAG 1256
 QY 401 AlaAlaArgAlaGlnGlyThrHisValLeuValHisCysLysMetGlyValSerArgSer 420
 DB 1257 GCTGCAGAGCAGACAGGACCCAGCTGCTGTCACCTCAAGATGGGGTTCAGCGCTCA 1316
 QY 421 AlaAlaThrValLeuAlaTyrAlaMetLysGlnTyrGlyCysSerLeuGlnAlaLeu 440
 DB 1317 GCGGCGCAGAGTCTGGCTATGCTCATGGAAGCATGACATGACCTGGAGCAGGCCCTG 1376
 QY 441 ArgHisValGlnGluLeuArgProLleAlaArgProAsnProGlyPheLeuArgGluLeu 460
 DB 1377 CGCCACGTGCAGAGACTCCGGGCCATCCCGGCCCAACCTGGGCTTCTCGGCACTG 1436
 QY 461 GlnLleTyrGlnGlyLleLeuThrAla 469
 DB 1437 CAGATCTACAGGCGATCTGACGAGCC 1463

RESULT 7
 ABQ73251
 ID ABQ73251 standard; cDNA: 2540 BP.
 AC ABQ73251:
 XX 30-SEP-2002 (first entry)
 XX

DE Human MAP kinase phosphatase splice form 3 cDNA sequence SEQ ID NO:3.
XX
KW Human: phosphatase; mitogen activated protein kinase phosphatase;
KW MAP kinase; enzyme; chromosome 11; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 5'UTR 1..87
FT CDS /*tag= a
FT 88..1314
FT /*tag= b
FT /*product= "MAP kinase phosphatase splice form 3"
FT 3'UTR 1315..2540
FT /*tag= c
XX
XX W0200242436-A2.
XX
XX 30-MAY-2002.
XX
XX 07-NOV-2001: 2001WO-US42995.
XX
XX 20-NOV-2000: 2000US-0715177.
XX 18-JAN-2001: 2001US-0761640.
XX
XX (PEKE) PE CORP NY.
XX
XX Wei M, Ketchum KA, Di Francesco V, Beasley EM;
XX WPI: 2002-575237/61.
XX P-PSDB: ABP51655.
XX
PT Novel isolated human phosphatase peptide useful for treating disorder
PT characterized by absence of, inappropriate or unwanted expression of
PT the phosphatase protein, and as immunogens to raise antibodies
XX
XX Claim 1; Fig 1C; 85pp; English.
XX
XX The present invention describes an isolated human phosphatase peptide
CC (1). (1) can be used for identifying a modulator of (1) by contacting
CC (1) with an agent and determining if the agent has modulated the
CC function or activity of (1). (1) is useful for identifying an agent that
CC binds to (1), by contacting (1) with an agent and assaying the contacted
CC mixture to determine whether a complex is formed with the agent bound
CC (1). The human phosphatases from the present invention are mitogen
CC activated protein (MAP) Kinase phosphatases. These human MAP kinase
CC phosphatases are located on chromosome 11. (1) and the polynucleotide
CC sequences encoding (1) can be used in gene therapy. The present sequence
CC encodes human MAP kinase phosphatase splice form 3 from the present
CC invention.
XX
XX Sequence 2540 BP; 552 A; 809 C; 739 G; 440 T; 0 other:
SQ
Alignment Scores:
Pred. No.: 1,35e-163 Length: 2540
Score: 2061.50 Matches: 408
Percent Similarity: 86.62% Conservative: 0
Best Local Similarity: 86.62% Mismatches: 0
Query Match: 84.63% Indels: 63
DB: 24 Gaps: 1
US-09-761-640-4 (1-471) x ABQ73251 (1-2540)
QY 1 MetalaleuValThrValSerArgSerProGlySerGlyAlaSerThrProValGly 20
DB 88 ATGGCCCTGGTCAAGTAGCCGTTCCGCCCGGCGACGGCGCTCCACGCCCTGGGG 147
QY 21 ProTrpAspGlnAlaValGlnArgArgSerArgLeuGlnArgArgGlnSerPheAlaVal 40
DB 148 CCTGGGACCGAGGGGCTCCACGAGAGTGACCTCCAGCAGGACAGACTTTGGGTG 207
QY 41 LeuArgGlyAlaValLeuGlyLeuGlnAspGlyGlyAspAsnAspAlaAlaGlnAla 60
|||||

DB 208 CTCCTGGGGCTGTCCTGGACTGCAGATGAGGGGACAAATGATGATCAGCAGAGGCC 267
QY 61 SerSerGluProThrGlnLysAlaProSerGluGluLeuHisGlyAspGlnThrAsp 80
DB 268 AGTTCTGAGCCACAGAGAGAGGCCCGGAGTGCAGCAGACGCTCCAGGGAGCAGACAG 327
QY 81 PheGlyGlnGlySerGlnSerProGlnLysGlnGluGlnArgGlnHisLeuHisLeu 100
DB 328 TTCCGGGCAAGAGATCCACAGATGCCAGAGAGGAGAGAGGACACCTGCACCTC 387
QY 101 MetValGlnLeuLeuArgProGlnAspAspIleArgLeuAlaGlnLeuGlnAlaPro 120
DB 388 ATGGTACAGCTGCTGAGGCCGACGATGATACCCCTGGGACGCCACCTGGAGGACCC 447
QY 121 ArgProProArgLeuArgTyrLeuLeuValSerThrArgGluGluGlyLeuSer 140
DB 448 CGGCTCCCGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 507
QY 141 GlnAspGluThrValLeuLeuGlnGlyAlaAspPheProAspSerSerProSerCysThr 160
DB 508 CAGATGAGAGCGGTCTCTCTGGCGCTGGATTTCCCTGACAGAGCTCCGCCAGCTGCAC 567
QY 161 LeuGlyLeuValLeuProLeuThrSerAspThrGlnValTyrLeuAspGlyAspGly 180
DB 568 CTGGGCTGTGCTTCCCTCCCTGAGTGACACCGAGTGTACTTGTAGATGAGAGCGGG 627
QY 181 PheSerValThrSerGlyGlyGlnSerArgIlePheLysProIleSerIleGlnThrMet 200
DB 628 TTCACCGTACGCTGTGGGCAAGCCGAGTTCACGCCATTCATCCAGCAGACATG 687
QY 201 TrpAlaThrLeuGlnValLeuHisGlnAlaCysGlnAlaAlaLeuGlySerGlyLeuVal 220
DB 688 TGG----- 690
QY 221 ProGlyGlySerAlaLeuThrTrpAlaSerHisTyrGlnGluArgLeuAsnSerGluGln 240
DB 690 ----- 690
QY 241 SerCysLeuAsnGluTrpThrAlaMetAlaAspLeuGluSerLeuArgProProSerAla 260
DB 690 ----- 690
QY 261 GluProGlyGlySerSerGluGlnGluGlnMetGluGlnAlaIleArgAlaGluLeuTrp 280
DB 691 -----TCTTCAGACAGAGCAGACATGAGCGCATCGTGTGAGCTGTGG 738
QY 281 LysValLeuAspValSerAspLeuGluSerValThrSerLysGluIleArgGlnAlaLeu 300
DB 739 AAGGTGTGATGATGACGTGAGTGTGAGAGTGTCACTTCCAAAGATCCGCCAGGCTGTG 798
QY 301 GluLeuArgLeuGlyLeuProLeuGlnGlnTyrArgAspPheIleAspAsnGlnMetLeu 320
DB 799 GAGCTGCGCCCTGGGGCTCCCTCCAGAGTACCGTGATCATGACACACAGATGCTG 858
QY 321 LeuLeuValAlaGlnArgAspArgAlaSerArgIlePheProHisLeuTyrLeuLysSer 340
DB 859 CTGCTGTGTGGCACACGGGACCGAGCTCCCGCATCTTCCCGACCTTACTGCGGCTCA 918
QY 341 GluTrpAsnAlaAlaAsnLeuGlnGluLeuGlnArgAsnArgValThrHisIleLeuAsn 360
DB 919 GAGTGAAGCGCAGCAACCTGGAGAGCTGCAGAGAACAGGAGTCCACCACTCTTGAAC 978
QY 361 MetAlaArgGluIleAspAsnPheTyrProGluArgPheThrTyrHisAsnValArgLeu 380
DB 979 ATGGCCCGGAGATTGAACAATCTTACCCCTGAGCGCTTACCTACCAACAATCTGCGCTC 1038
QY 381 TrpAspGluGluSerAlaGlnLeuLeuProHisTrpLysGluThrHisArgPheIleGln 400
DB 1039 TGGGATGAGAGTCCGCCAGCTGCTGCCCACTGGAAGAGAGCAGCCGCTTCAATTAG 1098
QY 401 AlaAlaArgAlaGlnGlyThrHisValLeuValHisCysLysMetGlyValSerArgSer 420
DB 1099 GCTGCAGAGACACAGGGGACCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1158

QY 421 A1aA1aThrValIleuAlaTyrAlaMetLysGlnTyrGluCysSerLeuGluGlnAlaLeu 440
 CC |||||||
 DB 1159 GGGGCGACAGCTGTGGCTATGCGATGAAAGCACTACGATGACGCTGGAGCGGCCCTG 1218
 CC |||||||
 QY 441 ArgHisValGlnGluLeuArgProIleAlaArgProAsnProGlyPheLeuArgGlnLeu 460
 CC |||||||
 DB 1219 CCGCAGCTGACAGAGCTCCGGCGCATGCGCGCCGACACCCCTGCTTCGCGCCAGCTG 1278
 CC |||||||
 QY 461 GlnIleTyrGlnGlyIleLeuThrAlaArgThr 471
 DB 1279 CAGATCTACCAAGGCGCTGACGCGCAAGAAC 1311
 CC |||||||
 RESULT 8
 ABL40805
 ID ABL40805 standard; cDNA; 2322 BP.
 AC ABL40805;
 XX
 DT 03-JUL-2002 (first entry)
 XX
 DE Human MAP kinase phosphatase-like enzyme encoding cDNA.
 XX
 KW Mitogen activated protein; MAP; MAP kinase phosphatase-like enzyme;
 KW antiasthmatic; antidiabetic; anorectic; cytostatic; cardiant; human;
 KW antiparkinsonian; cerebroprotective; neuroprotective; nootropic; gene;
 KW neuroleptic; anticonvulsant; anti-HIV; antirhythmic; hypotensive;
 KW antiallergic; dermatological; vulnerary; gene therapy; ss.
 OS Homo sapiens.
 XX
 Key Location/Qualifiers
 FT CDS 1..2322
 FT /tag= a
 FT /product= "MAP kinase phosphatase-like enzyme"
 FT /note= "contains internal codon deletions"
 FT /transl_except= "(pos: 156..157, aa: Leu)"
 FT /note= "there is an apparent one codon deletion which
 FT alters the reading frame"
 FT /transl_except= "(pos: 180..181, aa: Arg)"
 FT /note= "there is an apparent one codon deletion which
 FT alters the reading frame"
 FT /transl_except= "(pos: 414..415, aa: Val)"
 FT /note= "there is an apparent one codon deletion which
 FT alters the reading frame"
 FT /transl_except= "(pos: 1197..1198, aa: Tyr)"
 FT /note= "there is an apparent one codon deletion which
 FT alters the reading frame"
 FT /transl_except= "(pos: 2088..2089, aa: Xaa)"
 FT /note= "there is an apparent one codon deletion which
 FT alters the reading frame"
 XX
 PN W0200220732-A2.
 XX
 PD 14-MAR-2002.
 XX
 PE 27-AUG-2001; 2001WO-EP09848.
 XX
 PR 07-SEP-2000; 2000US-230709P.
 XX
 PA (FARB) BAYER AG.
 XX
 PI Liou J;
 XX
 DR WPI: 2002-339802/37.
 XX
 DR P-PSDB: ABB07845.
 XX
 PT New human mitogen activated protein kinase phosphatase-like enzyme
 PT polypeptide, regulators of which are useful for preventing, treating
 PT allergies including asthma, diabetes, obesity, cancer and
 PT cardiovascular diseases
 XX
 PS Claim 1; Fig 13; 134pp; English.

XX The invention relates to a purified human mitogen activated protein (MAP)
 CC kinase phosphatase-like enzyme polypeptide. The enzyme can be expressed
 CC by standard recombinant methodology. The MAP kinase phosphatase-like
 CC enzyme and encoding polynucleotides are useful for screening for
 CC modulators which are used for treating a MAP kinase phosphatase-like
 CC enzyme dysfunction related disease, such as asthma, a central nervous
 CC system disorder, diabetes, obesity, chronic obstructive pulmonary
 CC disease, cancer or a cardiovascular disease. The enzyme can be regulated
 CC to treat allergies including asthma, allergic rhinitis, atopic
 CC dermatitis, and anaphylaxis, central nervous system disorders such as
 CC brain injuries, Parkinson's disease, dementia, multiple sclerosis,
 CC stroke, Alzheimer's disease, Huntington's disease, schizophrenia, Pick's
 CC disease, Creutzfeldt-Jacob dementia, progressive nuclear palsy, and human
 CC immunodeficiency virus (HIV) dementia, and cardiovascular diseases
 CC including myocardial infarction, ischemic diseases of the heart, atrial
 CC and ventricular arrhythmia, hypertensive vascular diseases and peripheral
 CC vascular diseases. The enzyme is useful in diagnostic assays for
 CC detecting diseases and abnormalities or susceptibility to diseases or
 CC abnormalities related to the presence of mutations in the encoding
 CC nucleic acid sequences. The present sequence represents the human MAP
 CC kinase phosphatase-like enzyme polypeptide encoding cDNA.

SQ Sequence 2322 BP; 467 A; 751 C; 718 G; 386 T; 0 other;

Alignment Scores:

Prod. No.:	2e-157	Length:	2322
Score:	1987.50	Matches:	428
Percent Similarity:	86.63%	Conservative:	6
Best Local Similarity:	85.43%	Mismatches:	27
Query Match:	81.59%	Indels:	42
DB:	24	Gaps:	8

US-09-761-640-4 (1-471) x ABL40805 (1-2322)

QY 4 ValThrValSerArgSerProPro----- 11
 CC |||||||
 DB 53 GNCACAGTGAGCGGTGCGCCCGGTGGCAGCTGGCGCTCCACCGCCGCTGGGGCCCTG 112
 CC |||||||
 QY 12 -----GlySerGlyAlaSerThrProValGlyProThrPasp----- 23
 CC ||| |||
 DB 113 GAATTCCTGAGAGAGGGGAGGACAGCCCTCCCGCCCTCCACCTCGGGG-CTGCTCTCT 171
 CC ||| |||
 QY 24 -----GlnAlaValGlnArgSerArgLeuGlnArgGlnSerhealValIleu 41
 CC |||||||
 DB 172 CGGCAGAGCTGCTGCTCCAGGAGAGCTCCAGCTCCAGCAGAGA-----GCTTGG 219
 CC |||||||
 QY 42 Arg-----GlyAlaValLeuGlnLeuGlnAspGlyGlyAspAsnAspAlaAla 58
 CC ||| |||
 DB 220 CGGTGCTCTGCGGGCTGCTCGGAGCTGCAGAGATGAGAGGGGACAAATGATATCGACGA 279
 CC ||| |||
 QY 59 GluAlaSerSerGluProThrGlnLysAlaProSerGlnGlnLeuHis---GlyAsp 77
 CC |||||||
 DB 280 GAGGCCAGTTCGAGCCACAGAGAGGCCCGGAGTGAGAGAGACTCCACCGGGGAGACC 339
 CC |||||||
 QY 78 GlnThrAsp-PheGly-GlnGlySerGlnSerProGlnGlnGlnGlnArgGlnH 97
 CC |||||||
 DB 340 CAGACAGACTTTCGCTGCGATCCAGAGATCCCCAGAGAGGAGGAGGAGGAGCAGC 399
 CC |||||||
 QY 97 IsLeuHisLeuMetValGlnLeuLeuArgProGln-AspAsp11LeuValAlaAla--- 115
 CC ||| |||
 DB 400 AACTGCAACC-TCATGAGGCGTGTGAGGCCCGCAGAGATGACATCGACTTGGAGGCCA 458
 CC ||| |||
 QY 116 GlnLeuGlnAlaProArgProProArg-LeuArgTyr---LeuLeuValValSerThrAr 134
 CC ||| |||
 DB 459 ACCTGAGGACACCCCGGGCTCCCGGGTCCGAGTACCTTCGTGAGTTCACACG 518
 CC ||| |||
 QY 134 glnu-glyGlnGlyLeuSer-glnAspGlnThrValAlleuGly-ValAspPheProAs 153
 CC ||| |||
 DB 519 AGAAGGAGAAAGTGTGAGCCAGAGATGAGACGCTCTCTGCGACGTGGATTCTCCCTGA 578
 CC ||| |||
 QY 153 pSerSerSerProSerCysThrLeuGlyLeuValLeuProLeuThrPaspThrGlnVa 173
 CC ||| |||

```
Dh 579 CAGCAGCTCCCCAGCTGCACCTGGCTGCTTGGCCCTCTGGAGTGACACCCAGT 638
Oy 173 ILYRLSPGLYASPGLYGlyPheSerValThrSerGlyGlyInserArgIlePheLy 193
Dh 639 GTACTTATATGAGACGGGGGCTTACGCTGAGTCTGTGGGCAACCGGATCTTCA 698
Oy 193 sProIleSerIleGlnThrMetTrpAlaThrLeuGlnValLeuHisGlnAlaCysGluAl 213
Dh 699 GCCCATCTCCATCCAGACATGTGGGCACACTCCAGGATATTGCACCAAGCATGTAGAGC 758
Oy 213 aAlaLeuGlySerGlyLeuValProGlyGlySerAlaLeuThrTrpAlaSerHisTrpG 233
Dh 759 AGCTTACGACGGGCTTGTACCGGTGGCAGTGCCCTCACCTGGCCAGCCTACCA 818
Oy 233 nGluArgLeuSerGlnGlnInserCysLeuAsnGlnTrpThrAlaMetAlaAspLeuG 253
Dh 819 GGAGGAGCTGAACTCCGACAGAGCTGCTCATAGTGGAGGCTATGGCCGACTGGA 878
Oy 253 uSerLeuArgProProSerAlaGluProGlyGlySerSerGlnGlnGlnMetGluG 273
Dh 879 GTCTCTGGGGCTCCCGACGGCCGAGCTGGGGGCTCTCAGAACAGAGCATGTAGAGC 938
Oy 273 nAlaIleArgAlaGluLeuTrpLysValLeuAspValSerAspLeuGlnSerValThrSe 293
Dh 939 GCGGATCCGCTGCTGAGCTGTGGAAGTGTGGATGCTCGTACCTGGAGAGTGTGCTTC 998
Oy 293 rLysGluIleArgGlnAlaLeuGlnLeuArgGlnGlyLeuProLeuGlnGlnTrpArgAs 313
Dh 999 CAGAGAGATCCGCGAGGCTGTGGAGCTGCGCTGGGGCTCCCGCCCTCCAGCATACCGTA 1058
Oy 313 pPheIleAspAsnGlnMetLeuLeuValAlaGlnArgAspArgAlaSerArgIlePh 333
Dh 1059 CTTCATCGACAAACGAGATGCTGCTGTGTGGACAGGGGAGCGAGCTCCCGCATCTT 1118
Oy 333 eProHisLeuTrpLeuGlySerGlnTrpAsnAlaAlaAsnLeuGlnGlnLeuGlnArgAs 353
Dh 1119 CCCCCACCTTACCTGAGGCTGAGACTGAGACGACCAACCTGGAGGAGCTGCAGAGGA 1178
Oy 353 nArgValThrHisIleLeuAsnMetAlaArgGlnIleAspAsnPheTrpProGluArgPh 373
Dh 1179 CAGGGTACCCACATCTTG---ATGGCCCCGGAGATTCACAACTTCTACCTGAGCGCTT 1235
Oy 373 eThrTrpHisAsnValArgLeuTrpAspGluGlnuSerAlaGlnLeuLeuProHisTrpLy 393
Dh 1236 CACCTACACACATGTGGCCCTGTGGATGAGGAGTCCGCCACGCTGCGCGCATGTGGA 1295
Oy 393 sGluThrHisArgPheIleGluAlaAlaArgAlaGlnGlyThrHisValLeuValHisCy 413
Dh 1296 GGAGACGACCGCTTCATTGAGGCTGCAAGAGCACAGGGCACCCACGTGCTGGTCCACTG 1355
Oy 413 sLysMetGlyValSerArgSerAlaAlaThrValLeuAlaTrpAlaMetLysGlnTrpG 433
Dh 1356 CAAAGTGGGGCTGACGCCCTTACGCGCACAGTGTGCTGAGTATGCAATGAAAGACATAGA 1415
Oy 433 uCysSerLeuGlnGlnAlaLeuArgHisValGlnLeuLeuArgProIleAlaArgProAs 453
Dh 1416 ATGCAAGCTGGAGAGGGCCGTGCGGCACAGTGGAGAGAGCTCCGGCCCAATGCGCCCA 1475
Oy 453 nProGlyPheLeuArgGlnLeuGlnIleTrpGlnGlyIleLeuThrAla 469
Dh 1476 CCTGTGGCTTCTGCGCCAGCTGCAGATCTACAGAGGCATCTGAGGGCC 1524

RESULT 9
ID ABRN59832 standard; cDNA: 2061 BP.
AC ABRN59832;
XX 28-JUN-2002 (first entry)
DT Novel human coding sequence SFG ID NO: 243.
DE Human; antianaemic; vulnerability; antiinflammatory; immunomodulator;
KW Human; antianaemic; vulnerability; antiinflammatory; immunomodulator;
```

```
KW antinfertility; cerebroprotective; cytostatic; rheumatic; gene therapy;
KW neuroprotective; antiparkinsonian; protein therapy; EST;
KW expressed sequence tag; gene; ss.
OS Homo sapiens.
PN W0200222660-A2.
XX 21-MAR-2002.
PD 10-SEP-2001; 2001WO-US26015.
XX 11-SEP-2000; 2000US-0659671.
XX (HYSE-) HYSEQ INC.
PA Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX WPI: 2002-292408/33.
DR P-PSDB: ABB97419.
XX An isolated polynucleotide for treating diseases associated with its
PT encoded polypeptide such as cancer and multiple sclerosis -
PS Claim 1; SEQ ID NO 243; 509pp; English.
XX The present invention provides the protein and coding sequences of 444
CC novel human proteins. These were isolated from expressed sequence tags
CC (ESTs). They can be used to stimulate cell growth, to regulate
CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
CC e.g. in burn treatment, to regulate the immune system e.g. to treat
CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
CC infertility, to regulate haemostasis or thrombolysis e.g. to treat
CC stroke and cancer, to screen for drugs, to treat inflammatory conditions
CC e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.
CC Parkinson's disease. The present sequence is a coding sequence of the
CC invention.
XX
SQ Sequence 2061 BP; 415 A; 672 C; 605 G; 369 T; 0 other:
Alignment Scores:
Pred. No.: 5,12e-114 Length: 2061
Score: 1470.50 Matches: 290
Percent Similarity: 94.87% Conservative: 6
Best Local Similarity: 92.95% Mismatches: 15
Query Match: 60.37% Indels: 1
DB: 24 Gaps: 1
US-09-761-640-4 (1-471) x ABRN59832 (1-2061)
Oy 1 MetAlaLeuValThrValSerArgSerProProGlySerGlyAlaSerThrProValGly 20
Dh 147 ATGGCCCTGTGTCACAGTACGAGCGTTCGCCCGGGCACGGGCGCTCCACGCCCTGGGG 206
Oy 21 ProTrpAspGlnAlaValGlnArgArgSerArgLeuGlnArgArgGlnSerPheAlaVal 40
Dh 207 CCTTGGACACGCGCTCCAGGAGAGTGCACCTCCAGCAAGCAAGACAGCTTTCGGGTG 266
Oy 41 LeuArgGlyAlaValLeuGlyLeuGlnAspGlyGlyAspAsnAspAlaAlaGluAla 60
Dh 267 CTCCGTGGGGGCTGCTCTGGGACTGCAGATGGAGGAGCAATGATGATGACAGCAAGGCC 326
Oy 61 SerSerGluProThrGlnLysAlaProSerGlnGlnGlnLeuHisGlyAspGlnThrAsp 80
Dh 327 AGTTCTGAGCCCAACAGAGAGGCCCGGAGTGAAGAGAGCTCCAGGGGAGCCAGCAACAC 386
Oy 81 PheGlyGlnGlySerGlnSerProGlnLysGlnGlnGlnGlnArgGlnHisLeuHisLeu 100
Dh 387 TTGCGGCAAGGATCCCAAGATCCCAAGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 446
Oy 101 MetValGlnLeuLeuArgProGlnAspAspIleArgLeuAlaAlaGlnLeuGlnAlaPro 120
```

```
Dh 447 ATGTCACAGCTGCTGAGGCCGACGATGACATCCGCTGGCAGCCGACGTCGAGCACCC 506
Qy 121 ArgProProArgLeuArgTyrLeuLeuValValSerThrArgIuGluGlyLeuSer 140
Dh 507 CGGCTCCCGGGCTCCGCTACCTGCTGCTGCTTCTTACAGAGAGAGAGAGAGAGTCTGAGC 566
Qy 141 GlnAspGluThrValLeuLeuGlyValAspPheProAspSerSerSerProSerCysThr 160
Dh 567 CAGAGTAGAGACGGCTCCCTGGGGCTGGATTTCCCTGACAGCAGCTCCCGACCTGACCC 626
Qy 161 LeuGlyLeuValLeuProLeuThrSerAspPheGlnValTyrLeuAspGlyAspGly 180
Dh 627 CTGGGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 686
Qy 181 PheSerValThrSerGlyGlnSerArgIlePheLeuProIleSerIleGlnThrMet 200
Dh 687 TTCAGCGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 746
Qy 201 ThrPalaThrLeuGlnValLeuHisGlnAlaCysGluAlaAlaLeuGlySerGlyLeuVal 220
Dh 747 TGGGCGACACTCCAGCTATTTGCACCAAGCATGTGAGGACAGCTCTAGGACGGGCTTGT 806
Qy 221 ProGlyGlySerAlaLeuThrTrpPalaSerHisTyrGlnIuArgLeuAsnSerGluGln 240
Dh 807 CCGGCTGGCAGTGCCTCTACCTGGCCACCTACCAAGAGAGACTGAACTCCGAACAG 866
Qy 241 SerCysLeuAsnGluTrpThrAlaMetAlaAspLeuGluSerLeuArgProProSerAla 260
Dh 867 ACCTGCTCAATGAGTGGAGCGGCTATGGCGCACTGCACTCTGCGGCTCCACAGCC 926
Qy 261 GluProGlyGlySerSerLeuGlnGluGlnMetGluAlaIleArgAlaGluLeuTrp 280
Dh 927 GAGCTGGGGGCTCTCCACAGAGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 986
Qy 281 LysValLeuAspVal---SerAspLeuGluSerValThrSerGlyLeuArgGlnAla 299
Dh 987 AAAGGTGGAGCTGAGAGAGACCTCAGAGAGACCTGACATGCCAGAGGTCTTCTCTCC 1046
Qy 300 LeuGluLeuArgLeuGlyLeuProLeuGlnGlnTyr 311
Dh 1047 CACGAGTCTTCATGATGAAGAGCTCTGCGAGCCCTTC 1082

RESULT 10
AAD09491
ID AAD09491 standard; DNA: 6374 BP.
XX
AC AAD09491;
XX
DT 10-SEP-2001 (first entry)
XX
DE Human SGP006 phosphatase polypeptide encoding DNA.
XX
KW Human; SGP006 phosphatase polypeptide; phosphatase-related disease;
KW immune-related disorder; ocular disease; organ transplant rejection;
KW infection; diabetes; pain; sexual dysfunction; Alzheimer's disease;
KW metabolic disorder; haematopoietic cancer; mood disorder; cardiant;
KW Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis;
KW cardiovascular disease; brain; neuronal-associated disease; dyskinesia;
KW attention disorder; cognition disorder; psychotic disorder; cytoskeletal
KW neurological disorder; virucide; nootropic; cerebroprotective; therapy;
KW neuroprotective; antibacterial; vulnery; tranquiliser; antiasthmatic;
KW hypotensive; immunosuppressive; antipsoriatic; analgesic; hypertensive;
KW antifungal; dual specificity phosphatase; DSP; MAP kinase phosphatase;
KW MKP; mdgraine; chromosome 12q21.3-q22; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 34..3183
FT /*tag= a
FT /product= "Human SGP006 phosphatase polypeptide"
XX
PN MO200146394-A2.
```

```
XX
PD 28-JUN-2001.
XX
PF 21-DEC-2000; 2000MO-US34736.
XX
PR 21-DEC-1999; 99US-0173255.
PR 28-DEC-1999; 99US-0175766.
PR 25-JAN-2000; 2000US-0178078.
PR 31-JUN-2000; 2000US-0179301.
XX
PA (SUGEN-) SUGEN INC.
XX
PI Plowman GD, Martinez R, Whyte D, Manning G, Sudarsanam S, Hill RJ;
PI Flanagan P;
XX
DR WPI; 2001-418058/44.
XX
P-PSDB; AAE04833.
XX
PT Novel phosphatase polypeptide useful for treating cancers,
PT immune-related diseases and disorders, cardiovascular disease, brain or
PT neuronal-associated diseases and metabolic disorders
XX
PS Claim 29; Fig 1; 186pp; English.
XX
CC The present invention relates to phosphatase polypeptides, nucleotide
CC sequences encoding them, as well as various products and methods useful
CC for the diagnosis and treatment of various phosphatase-related diseases
CC and conditions. Substance that modulates the activity of phosphatase
CC polypeptide is used to treat immune-related diseases and disorders,
CC cardiovascular disease, brain or neuronal-associated diseases and
CC metabolic disorders, including cancers of tissues, cancers of
CC haematopoietic origin, diseases of central and peripheral nervous
CC system, Alzheimer's disease, Parkinson's disease, multiple sclerosis,
CC amyotrophic lateral sclerosis, viral infections, infections caused by
CC prions, bacteria and fungi, ocular diseases, diabetes, migraines, pain,
CC sexual dysfunction, mood disorders, attention disorders, cognition
CC disorders, hypotension, hypertension, psychotic disorders, neurological
CC disorders, dyskinesias and organ transplant rejection. The present
CC sequence is a DNA encoding human SGP006 phosphatase polypeptide. This
CC sequence is classified as dual specificity phosphatase (DSP) and MAP
CC kinase phosphatase (MKP). SGP006 gene maps to chromosomal position
CC 12q21.3-q22.
XX
SQ Sequence 6374 BP; 1604 A; 1674 C; 1659 G; 1437 T; 0 other:
XX
XX
Alignment Scores:
Pred. No.: 2,65e-83 Length: 6374
Score: 1113.00 Matches: 232
Percent Similarity: 65.14% Conservative: 80
Best Local Similarity: 48.43% Mismatches: 127
Query Match: 45.69% Indels: 40
DB: 22 Gaps: 8
XX
US-09-761-640-4 (1-471) x AAD09491 (1-6374)
Qy 1 MetalAlaLeuValThrValSerArgSerProGlySerGlyAlaSerThrProValGly 20
Dh 34 AARGGCGCTGAGTACCTCGAGGCTCGCCACAGCCAGCGCGCTCTCTCGGCCAGC 93
Qy 21 ProTrpAsp-----GlnAlaValGlnArgArgSerArgLeuArgGln 36
Dh 94 AACACGAGGTTGGAGCTGCGCGACGAGAGATGCAAAATTAAACCTGAGCTTAGTGAG 153
Qy 37 SerPheAlaValLeuArgIleValValLeuGlyLeuGlnAspGlyGlyAspAsnAsp 56
Dh 154 AGCTTTTCATGAGTGAAGAGCGACGCTCTTCTTACCAACAGGGA----- 198
Qy 57 AlaAlaGlnAlaSerSerGluProThrGluLysAlaProSerGluGluLeuHisGly 76
Dh 199 -----AGCAGCGCCTCA----- 210
Qy 77 AspGlnThrAspPheGlyGlyGlySer---GlnSerProGlnLysGlnGlu---GlnGln 94
```

```

Db 211 -----GGCCAGCGAGTCTTCAGACCCCCACAGCATCAGTGATCTG 255
QY 95 ArgGlnHisLeuHisLeuMetValGlnLeuLeuArgProGlnAspIleArgLeuAla 114
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 256 CCTCAACATCTTCAGATGATGATCAACCTTCTGCGTTGGAAGACAGATCAAGCTGGGA 315
QY 115 AlaGlnLeuGlnAlaProArgProProArgLeuArgTyrLeuLeuValAla -----Ser 132
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 316 GTGCCCTGTGAGAGCGCGCTGGCGGACCGGTCGGGTACATGGTGTGTACAGCAC 375
QY 133 ThrArgGlnGlnGlnGlnLeuSerGlnAspGlnThrValLeuLeuGlnValAlaSerPhePro 152
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 376 GGGCGCCAG-----GACACCGAGGAGAAATATCTTCTGCGAGTGCATCTTTC 423
QY 153 AspSerSerSerProSerCysThrLeuGlnGlnValLeuProLeuThrSerAspThrGln 172
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 424 ACTAAGAAAGTAAAGCTGACACCATGGATGTTCTCCAGCTGGGAGCACACGAAA 483
QY 173 ValTyrLeuAspGlyAspGlyGlyPheSerValThrSerGlyGlyGlnSerArgIlePhe 192
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 484 ATCCACCTGATGAGAGATGATGGTTCAGCGTACAGACAGACAGAAAGCATGACATATT 543
QY 193 LysProIleSerIleGlnThrMetThrAlaThrLeuGlnValLeuHisGlnAlaCysGln 212
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 544 AAGCTGTCTGTCTGACAGGCCATGTGTCTGCTGCCCTGCAGGTCTTACACAGGCTGCGAA 603
QY 213 AlaAlaLeuGlySerGlyLeuValProGlyGlySerAlaLeuThrThrPheSerHisTyr 232
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 604 GTGGCCCGGAGGCACACACTCTCCCGGGGTGATGCTCTATCTAGGCTACCTACTAT 663
QY 233 GlnGlnAlaTyrLeuAsnSerGlnGlnSerCysLeuAsnGlnTyrThrAlaMetAlaSerLeu 252
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 664 GAGACTGTCATGATCGCAGACGATGCATCAAGATGAGAACGATGACGAGACCTG 723
QY 253 GluSerLeuArgPro-----ProSerAlaGlnProGlyGlySerSerGlnGlnGln 270
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 724 GAGTCTACGGCGCCGACTCCCGCGCTATTGTGACACCCCACTGAAGCGAAAG 783
QY 271 MetGlnGlnAlaIleArgAlaGlnLeuTyrPlyValLeuAspValSerAspLeuGlnSer 290
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 784 ACCGAGCGCTCATCAAAAGCCAGCTCCGAGACATCATCATGACGACGATCTAGAAAAT 843
QY 291 ValThrSerLysGlnIleArgGlnAlaLeuGlnLeuArgLeuGlyLeuProLeuGlnGln 310
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 844 GTGACTTCCAAAGAGATTCGTATGATTAAGAGAAACAGATGAATTGACTTGAAGAA 903
QY 311 TyrArgAspPheIleAspAsnGlnMetLeuLeuValAlaGlnArgAspArgAlaSer 330
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 904 CTCAAGGAATTTATAGACAAATGAGATGCTACTTATCTTGGGACAGATGCAAGCCCTCC 963
QY 331 ArgIlePheProHisLeuTyrLeuGlySerGlnTyrPheAlaAlaAsnLeuGlnGln 350
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 964 CTTATCTTTCATCATCTTTATCTGCGCTCTGATGATGATGATCAATCTGAGGAGACTG 1023
QY 351 GlnArgAsnArgValThrHisIleLeuAsnMetAlaArgGlnIleAspAsnPheTyrPro 370
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1024 CAGGCTCGAGGGGTGATTAATTAATCTTACCAAGAAATGATATTTTCTCT 1083
QY 371 GlnArgPheThrThrHisAsnValArgLeuTyrPheGlnGlnSerAlaGlnLeuLeuPro 390
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1084 GCGTTATTTGCAATCAATCAATCCGAGTCTGATGAAGAACCAACAGACCTCTCCGCC 1143
QY 391 HisTyrPlyGlnThrHisArgPheIleGlnAlaAlaArgAlaGlnGlyThrHisValLeu 410
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1144 CACTGGAATGAAGCGTATCATTTTAAACAAGGAGAGAGAACCATTCAGAGTCCG 1203
QY 411 ValHisCysLysMetGlyValSerArgSerAlaAlaThrValLeuAlaTyrAlaMetLys 430
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1204 GTGCAATTCGAAAATGAGCGTGCAGTCCGCTCCACAGCATAGCTATGCAATGAAG 1263
QY 431 GlnTyrGlnCysSerLeuGlnAlaLeuArgHisValGlnGlnLeuArgProIleAla 450
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1264 GAATTCGGCTGGCTCTGAAAAGCATATTAATATGAAGCAGAAAGCGACATCAAG 1323

```

```

QY 451 ArgProAsnProGlyPheLeuArgGlnLeuGlnIleTyrGlnGlyIleLeuThrAla 469
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1324 CGCCCAACCGCGGCTTTATGAGCAGCTGTCTGAGTATGAAGCATTTTGATGCA 1380
RESULT 11
AAD09493
ID AAD09493 standard; DNA; 2260 BP.
XX
AC AAD09493;
XX
DE 10-SEP-2001 (first entry)
XX
DE Human SGP001 phosphatase polypeptide encoding DNA.
XX
KW Human; SGP001 phosphatase polypeptide; phosphatase-related disease;
KW Immune-related disorder; organ disease; organ transplant rejection;
KW Infection; diabetes; pain; sexual dysfunction; Alzheimer's disease;
KW Metabolic disorder; haematopoietic cancer; mood disorder; cardiac;
KW Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis;
KW Cardiovascular disease; brain; neuronal-associated disease; dyskinesia;
KW Attention disorder; cognition disorder; psychotic disorder; cytosolic;
KW Neuroprotective; antiviral; nootropic; cerebroprotective; therapy;
KW Neuroprotective; antibacterial; vulnerrary; tranquilliser; antiaesthetic;
KW Hypotensive; immunosuppressive; antipsoriatic; analgesic; hypertensive;
KW Antifungal; dual specificity phosphatase; DSP; MAP kinase phosphatase;
KW MKP; migraine; chromosome Xp11.1-11.3; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 709..2205
FT FT /*tag= a
FT FT /product= "Human SGP001 phosphatase polypeptide"
XX
PN WO200146394-A2.
XX
PD 28-JUN-2001.
XX
PE 21-DEC-2000; 2000MO-US34736.
XX
PR 21-DEC-1999; 99US-0173255.
PR 28-DEC-1999; 99US-0175766.
PR 25-JAN-2000; 2000US-0178078.
PR 31-JAN-2000; 2000US-0179301.
XX
PA (SUGEN-) SUGEN INC.
XX
PI Plowman GD, Martinez R, Whyte D, Manning G, Sudarsanam S, Hill RJ;
PI Flanagan P;
XX
DR WPI. 2001-418058/44.
XX
DR P-PSDB; AAE04835.
XX
PT Novel phosphatase polypeptide useful for treating cancers,
PT immune-related diseases and disorders, cardiovascular disease, brain or
PT neuronal-associated diseases and metabolic disorders
XX
PS Claim 29; Fig 1; 186pp; English.
XX
CC The present invention relates to phosphatase polypeptides, nucleotide
CC sequences encoding them, as well as various products and methods useful
CC for the diagnosis and treatment of various phosphatase-related diseases
CC and conditions. Substance that modulates the activity of phosphatase
CC polypeptide is used to treat immune-related diseases and disorders,
CC cardiovascular disease, brain or neuronal-associated diseases and
CC metabolic disorders, including cancers of tissues, cancers of
CC haematopoietic origin, diseases of central and peripheral nervous
CC system, Alzheimer's disease, Parkinson's disease, multiple sclerosis,
CC amyotrophic lateral sclerosis, viral infections, infections caused by
CC prions, bacteria and fungi, ocular diseases, diabetes, migraines, pain,
CC sexual dysfunction, mood disorders, attention disorders, cognition
CC disorders, hypotension, hypertension, psychotic disorders, neurological

```


PF 01-FEB-2001; 2001WO-US03429.
XX
PR 02-FEB-2000; 2000US-0179886.
XX
PA (CEPT-) CEPTIR INC.
XX
PI Lucie RM, Wei B;
XX
DR MPI: 2001-48887/53.
DR P-PSDB: AAE06775.
XX
PT New isolated dual-specificity phosphatase polypeptide for treating
PT cancer, graft-versus-host disease, autoimmune diseases, allergies,
PT metabolic diseases, abnormal cell growth and abnormal cell
PT proliferation -
XX
PS Claim 56; Fig 3; 81pp; English.
XX
CC The present sequence is human dual-specificity phosphatase (DSP)-13
CC cDNA. Inactivation of mitogen-activated protein kinase (MAP-kinase)
CC is mediated by dephosphorylation at a dual phosphorylation motif by DSP
CC which is referred to as MAP-kinase phosphatase. An agent that modulates
CC DSP is useful for treating a disorder selected from Duchenne muscular
CC dystrophy, cancer, graft-versus-host disease (GVHD), autoimmune diseases,
CC allergies, metabolic diseases, abnormal cell growth, abnormal cell
CC proliferation and cell cycle abnormalities. DSP is useful for
CC identifying antibodies and other agents that inhibit DSP-12 and/or
CC DSP-13 activity. DSP and the agents identified are useful for
CC modulating cell proliferation, differentiation and survival. DSP is
CC useful in screening assays for modulators of enzyme activity and
CC substrate binding and for dephosphorylating a substrate of DSP-12 or
CC DSP-13.
XX
SQ Sequence 1711 BP; 514 A; 396 C; 390 G; 411 T; 0 other;
XX
Alignment Scores:
Pred. No.: 2,42e-78 Length: 1711
Score: 1045.00 Matches: 200
Percent Similarity: 66.37% Conservative: 94
Best Local Similarity: 45.15% Mismatches: 107
Query Match: 42.90% Indels: 42
DB: Gaps: 4
US-09-761-640-4 (1-471) x AAD12966 (1-1711)
QY 36 GlnSerPheAlaValLeuArgGlyAlaValLeuGlyLeuGlnAspGlyGlyAspAsnsp 55
DB 240 GAGAGCTTCTTAAGTCTGCTCAAGGTCCTCCCTTTTCTTACACAGG----- 284
QY 56 AspAlaAlaGluAlaSerSerGlnProThrGlnLysAlaProSerGlnGlnLutLeuHis 75
DB 284 ----- 284
QY 76 GlyAspGlnThrAspPheGlyGlnGlySerGlnSerProGlnLysGlnGlnGln 94
DB 285 -----GGAATGCGCTCATCCACCAAGCAATGAGCCAGACGAGAGCAAC 326
QY 95 -----ArgGlnHisLeuHisLeuMetValGlnLeuLeuArgProGln 108
DB 327 AAGCATGACGAGGATCTCCACACAGCATCTCCACGAATGTCATTCTACCGCCAGAA 386
QY 109 AspAspIleArgLeuAlaGlnLeuGlnAlaProArgProProArgLeuArgTyrLeu 128
DB 387 GACACATCAGAGCTGCGCTGAAGACTGAAAGTACTTACCAAGATCGAAGACGCTATATG 446
QY 129 LeuValValSerThrArgGlnGlyGlnGlyLeuSerGlnAspGlnThrValLeuLeuGly 148
DB 447 GTAGTGTTTCAATCAATAGTAGACAA-----GACACTGAAGAAGCATCTCTTAGGA 500
QY 149 ValAspPheProSerSerSerProSerCysThrLeuGlnGlyLeuValLeuProLeuTyr 168
DB 501 ATGGATTCTCCTCTTAATGACAGTAGCACTGTGTACCATGGGCTTAGTTTGGCTCTCTGG 560

QY 169 SerAspThrGlnValTyrLeuAspGlyGlyPheSerValThrSerGlyGln 188
DB 561 AGCGACACGCTAATTCATTGATGATGGTGTGCTGAGTTCAGTATCGACGATTAACGA 620
QY 189 SerArgIlePheLysProIleSerIleGlnThrMetThrPalaThrLeuGlnValLeuHis 208
DB 621 GTTCACATATTCMAACCTGATCTGTGCGGCAATGTGCTGTGACACTACAGACTTACAC 680
QY 209 GlnAlaCysGlnAlaAlaLeuGlySerGlyLeuValProGlyGlySerAlaLeuThrTrp 228
DB 681 AAGCTTGTGAGTGTGCCAGACCATTAATCTACCCAGGACGCTATTTCTCACTTGG 740
QY 229 AlaSerHisTyrGlnGlnLysArgLeuAsnSerGlnSerCysLeuAsnGlnThrPala 248
DB 741 GTGAGTTTATTATGAGCCCATATCAACTCAGATCATCTCCTGATGATGATGATGATGCA 800
QY 249 MetAlaAspLeuGlnSerLeuArgPro-----ProSerAlaGlnProGlyLysSerSer 266
DB 801 ATGCAAGATGTACAGTCCACCGCCGACCTCTCCAGCTCTTCCACCGACATACCTACT 860
QY 267 GlnGlnGlnGlnMetGlnGlnAlaIleArgAlaGlnLeuTrpLysValLeuAspValSer 286
DB 861 CACGCTGAACGACAGCAAGAAAGCTAATTAAACCAATTTAAGGAGATCTGATCGAAG 920
QY 287 AspLeuGlnSerValThrSerLysGlnLysArgGlnAlaLeuGlnLeuArgLeuGlyLeu 306
DB 921 GATTGGAGAAATATTACATCCAAAGATTAAGACAGATGGAATGGAATGGTGTGC 980
QY 307 ProLeuGlnGlnTyrArgAspPheIleAspAsnGlnMetLeuLeuValAlaGlnArg 326
DB 981 AACTTGGCGGAATTCAGAAATTTATAGCAATGAAATGATGTATCTCTGTGTCAATAG 1040
QY 327 AspArgAlaSerArgIlePheProHisLeuTyrLeuGlySerGlnTrpAsnAlaIleAsn 346
DB 1041 GATAGCCCTACACAGATATTGATGATGCTGTCTGCGCTCGATGATGATGATGCTTCCAAC 1100
QY 347 LeuGlnGlnLeuGlnLysArgAsnArgValThrHisIleLeuAsnMetAlaArgLysLeuAsp 366
DB 1101 TTAGAGGACTTACAGAACGAGGAGTACGGTATATCTTAATGATCCTCGAGAGATAGAT 1160
QY 367 AsnPheTyrProGlnArgPheThrThrHisAsnValArgLeuTrpAspGlnLutSerAla 386
DB 1161 AACTTCTTCCCGAGGAGTCTTGTAGATATCAATTCGCGGTATATGATGAAAGACGCAACG 1220
QY 387 GlnLeuLeuProHisTyrLysGlnThrHisArgPheIleGlnAlaAlaArgAlaGlnGly 406
DB 1221 GATCTCTGCGCTTACGATGACACTTACAAATTCATCTTAAAGCAAGAAACATGGA 1280
QY 407 ThrHisValLeuValHisCysLysMetGlyValSerArgSerAlaAlaThrValLeuAla 426
DB 1281 TCTAATATGCTTGTGCACATCGCAAAATGGGGGTGAGTCCGCTCAGCCTCCACCGTATGCC 1340
QY 427 TyrAlaMetLysGlnTyrGlnCysSerLeuGlnGlnAlaLeuArgHisValGlnLutLeu 446
DB 1341 TATGCAATGMAAGATATGCTGTAATCTGGACCGGCTTATGATGTAAGAAAGA 1400
QY 447 ArgProIleAlaArgProAsnProGlyPheLeuArgGlnLeuGlnIleLysGlnGlyIle 466
DB 1401 CGAAGGATACCAAGCCCAACCAAGCTTATGAGACATGGAAGATGAGAGTACAGGGATC 1460
QY 467 LeuThrAla 469
DB 1461 TTGCTGGCA 1469
RESULT 13
AAH22966 standard; cDNA; 1771 BP.
ID AAH22966
AC AAH22966;
XX
XX
DT 17-SEP-2001 (first entry)
XX Human phosphatase (PP) encoding cDNA (clone ID 637331CB1).

XX Phosphatase; PP; Alzheimer's disease; Huntington's disease; human;
 KW immune disorder; cancer; neotrophic; immunomodulatory; cytostatic;
 KW neuroprotective; anticonvulsant; ss.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FT CDS 24..1478
 FT /tag= a
 FT /product= "phosphatase"
 FT
 XX MO200153469-A2.
 XX 26-JUL-2001.
 XX 18-JAN-2001; 2001WO-US02088.
 XX 21-JAN-2000; 2000US-0177719.
 XX 28-JAN-2000; 2000US-0178988.
 XX 25-FEB-2000; 2000US-0184959.
 XX 17-MAR-2000; 2000US-0190142.
 XX (INCY) INCYTE GENOMICS INC.
 XX Bandman O, Tang YT, Azimzai Y, Yue H, Baughn MR, Hillman JL,
 PI Lal P, Wang E, Gandhi AR, Policky JL, Mathur P;
 DR WPI: 2001-451905/48.
 DR P-PSDB: AAB85363.
 XX Novel phosphatases useful for treating diseases associated with
 PT decreased expression of functional phosphatases, e.g., Alzheimer's
 PT disease, Huntington's disease, immune disorders, and cancers -
 XX
 PS Claim 5: Page 102-103; 103pp; English.
 XX
 CC The invention provides human phosphatases (PP) and polynucleotides
 CC encoding the phosphatases. The polypeptides can be expressed by standard
 CC recombinant methodology. The PP are useful for treating a disease or
 CC condition associated with decreased expression of functional
 CC phosphatases. Compositions containing agonists or antagonists of PP may
 CC be used to treat a disease associated with decreased expression or
 CC overexpression of PP, respectively. Such diseases may include
 CC Alzheimer's disease, Huntington's disease, immune disorders, and
 CC cancers. The present sequence represents a human phosphatase encoding
 CC cDNA.
 CC
 SQ Sequence 1771 BP; 529 A; 395 C; 378 G; 469 T; 0 other;
 XX
 Alignment Scores:
 Pred. No.: 6.44e-75 Length: 1771
 Score: 1004.50 Matches: 199
 Percent Similarity: 64.04% Conservative: 93
 Best Local Similarity: 43.64% Mismatches: 119
 Query Match: 41.24% Indels: 45
 DB: 22 Gaps: 5
 US-09-761-640-4 (1-471) x AAH22966 (1-1771)
 QY 1 MetaLalaValIthValSerArgSerProGlySerGlyAlaSerThrProValGly 20
 DB 24 AAGGCTTGTGTGTCACGCGTCCAGCTACCTACCCGACGACGACCTCCGCGCC 83
 QY 21 PGTTPASGAlaValGlnAArgArgSerArgLeuGlnAArgArg-----GlnSer 37
 DB 84 TCGGAGGACGACAGTGGGAGGAAGATGCCGCTACAGCCGAGGAGCATCAGCGAGAGC 143
 QY 38 PheAlaValIleuArgGlyAlaValLeuGlnGlyLeuGlnAspGlyLysAspAsnAspAla 57
 DB 144 TTTCCTAAGCTGCAAGAGTGTGCTTTTCTTACACAG-----182
 QY 58 AlaGlnAlaSerSerGluProThrGlnLysAlaProSerGlnGlnGlnLysGlyAsp 77

DB 182 -----182
 QY 78 GlnThrAspPheGlyGlnGlySerGlnSerProGlnLysGlnGlnGln----- 94
 DB 183 -----GGAANTGGCTATCCACCAAGATCAGCCACAGAGCAACAGCAT 230
 QY 95 -----ArgGlnHisLeuHisLeuMetValGlnLeuLeuArgProGlnAsp 110
 DB 231 GCAGGCGATCTCCACAGCATCTCCACAGCATGTTCTTACTCGCGCCAGAGCAAC 290
 QY 111 IleArgLeuAlaAlaGlnLeuGlnAlaProArgProProArgLeuArgGlyLeuLeuVal 130
 DB 291 ATCAGCGCTGGCTTAAGAGTGAAGTACTTACCAAGATCGAAGCAGCTATGTGATGTG 350
 QY 131 ValSerThrArgGlnGlnGlyLeuSerGlnAspGlnThrValIleuLeuGlyAlaAsp 150
 DB 351 GTTTCACCTAATGATGAGCAA-----GACACTGAAGAAAGACATGCTCTAGGAATGAT 404
 QY 151 PheProAspSerSerSerProSerCysThrLeuGlyLeuValLeuProLeuThrSerAsp 170
 DB 405 TTCTCCTTAATGACAGTGAAGTGTACATGCGCTTGTGTTGGCTCTCGAGCGAC 464
 QY 171 ThrGlnValIleuLeuAspGlyLysGlyLysSerValThrSerGlyGlyLysAsnArg 190
 DB 465 ACGCTAATTATTTGGATGATGATGGTTCAGTGTATGACGATGATACAGAGTTTAC 524
 QY 191 IlePheLysProIleSerIleGlnThrMetThrPalaThrLeuGlnValLeuHisGlnAla 210
 DB 525 ATATTCAAACTGTATCTTGTGACGAGCATGTGCTGCTGCTACAGAGCTTACCAAGCT 584
 QY 211 CysGlnAlaAlaLeuLeuLysSerGlyLeuValProGlyGlySerAlaLeuThrTrpAlaSer 230
 DB 585 TGTGAAGTGGCCAGACGCGATACCTACTACCCAGGAGCGCTATTCTCTGCGGAGT 644
 QY 231 HistGlyGlnGlnArgLeuAsnSerGlnGlnSerGlyLeuAsnGlnThrPheThrAlaMetAla 250
 DB 645 TATTATGAGAGCCATATTCATCACTGATCAATCTCAGTCAATGAATGAATGAAATGCA 704
 QY 251 AsnLeuGlnSerLeuArgPro-----ProSerAlaGlnProGlyLysSerGlnGln 268
 DB 705 GATGTACACTCCACCGCGCCGACCTCTCCAGCTCTTACCGAGACTACTACTGACACT 764
 QY 269 GlnGlnMetGlnGlnAlaIleArgAlaGlnLeuTrpLysValLeuAspValSerAspLeu 288
 DB 765 GAACGACAGAAAGGCTATTAACCAATTAAGGAGATGATGTGCGAAGAGATTGG 824
 QY 289 GluSerValThrSerLysGlnIleArgGlnAlaLeuGlnLeuArgLeuGlyLeuProLeu 308
 DB 825 GAGATATTACATCCAAAGAGATAGACAGAGTTGGAATGCAATGGTGCAACTTG 884
 QY 309 GlnGlnTrpArgAspPheIleAspAsnGlnMetLeuLeuValAlaGlnArgAspArg 328
 DB 885 CGGGAATTCAAGGAATTTATAGCATGAATGATGATCTCTGTCGCAATGGATGAC 944
 QY 329 AlaSerArgIlePheProHisLeuTrpLeuGlySerGlnTrpAsnAlaAlaAsnLeuGln 348
 DB 945 CCTPACACAGATATTGTGAGCATGTCTCGGCTCGAAGATGGAATCTCCCAACTTAAGG 1004
 QY 349 GlnLeuGlnArgAsnArgValThrHisIleLeuAsnMetAlaArgGlnIleAspAsnPhe 368
 DB 1005 GACTTACAGAAACGAGGAGGATGATATCTGAAATGATCTCGAAGATAGATTAATTCTTC 1064
 QY 369 TyrProGlnArgPheThrTrpHisAsnValArgLeuTrpAspGlnGlnLeuSerAlaGlnLeu 388
 DB 1065 TTCCCAAGAGTCTTTGAGATATCATTAATTCGGGTATGATGATGAAGCAACGATCTTC 1124
 QY 389 LeuProHisTrpLysGlnThrHisArgPheIleGlnAlaAlaArgAlaGlnGlyThrHis 408
 DB 1125 CTGCGCTACTGCAATGACATCTTCAATTCATCTTAAGCAAAAGACATGATGTAAA 1184
 QY 409 ValLeuValHisCysLysMetGlyValSerArgSerAlaAlaThrValLeuAlaArgAla 428

DB	Accession	Gene	Protein	Length	Score	Similarity
Db	1185	TGCTTGACACCTGCAAAATGGGGGCGTAGTCGCTCAGCCCTCCACCGCTGATTGGCTATGCA	1244			
Oy	429	MetylsyntyrilgucysSerleuglucinalaleaRghisValcIn	444			
Db	1245	ATGAAGCATATGCTGGAATCTGACCGAGCCTCGACCGCAATTCGG	1292			
RESULT 14						
AAID12965						
ID	AAID12965	standard; cDNA; 1949 BP.				
XX	AC	AAID12965;				
XX	AD	16-OCT-2001 (first entry)				
DE	XX	Human dual-specificity phosphatase (DSP)-12 cDNA.				
XX	XX	Human: dual-specificity phosphatase; DSP-12: dual phosphorylation motif;				
KW	KW	mitogen-activated protein kinase; MAP-kinase; graft-versus-host disease;				
KW	KW	GvHD; Duchenne muscular dystrophy; therapy; cancer; autoimmune disease;				
KW	KW	allergy; metabolic disease; cell growth; cell proliferation; cytostatic;				
KW	KW	cell cycle abnormality; cell differentiation; anti-allergic; muscular;				
XX	OS	immunosuppressive; ss.				
XX	XX	Homo sapiens.				
XX	XX	Key	Location/Qualifiers			
FT	FT	CDS	95..1753			
FT	FT	/tag- a	/product= "Human dual-specificity phosphatase (DSP)-12"			
XX	XX	MO200157221-A2.				
XX	XX	09-AUG-2001.				
PD	PD	01-FEB-2001; 2001MO-US03429.				
XX	XX	02-FEB-2000; 2000US-0179886.				
PR	PR	(CEPT-) CEPTYR INC.				
PA	PA	Luche RM, Wei B;				
PI	PI	WP1: 2001-488887/53.				
XX	DR	P-PSDB; AAE06772.				
XX	XX	New isolated dual-specificity phosphatase polypeptide for treating				
PT	PT	cancer, graft-versus-host disease, autoimmune diseases, allergies,				
PT	PT	metabolic diseases, abnormal cell growth and abnormal cell				
XX	XX	proliferation -				
PS	PS	Claim 7; Fig 1; 81pp; English.				
XX	XX	The present sequence is human dual-specificity phosphatase (DSP)-12				
CC	CC	cDNA. Inactivation of mitogen-activated protein kinase (MAP-kinase)				
CC	CC	is mediated by dephosphorylation at a dual phosphorylation motif by DSP				
CC	CC	which is referred to as MAP-kinase phosphatase. An agent that modulates				
CC	CC	DSP is useful for treating a disorder selected from Duchenne muscular				
CC	CC	dystrophy, cancer, graft-versus-host disease (GVHD), autoimmune diseases				
CC	CC	allergies, metabolic diseases, abnormal cell growth, abnormal cell				
CC	CC	proliferation and cell cycle abnormalities. DSP is useful for				
CC	CC	identifying antibodies and other agents that inhibit DSP-12 and/or				
CC	CC	DSP-13 activity. DSP and the agents identified are useful for				
CC	CC	modulating cell proliferation, differentiation and survival. DSP is				
CC	CC	useful in screening assays for modulators of enzyme activity and				
CC	CC	substrate binding and for dephosphorylating a substrate of DSP-12 or				
CC	CC	DSP-13.				
XX	XX	Sequence 1949 BP; 516 A; 509 C; 513 G; 411 T; 0 other;				
XX	XX	Alignment Scores:				
XX	XX	Pred. No.:	8.4e-73	Length:	1949	
XX	XX	Score:	980.00	Matches:	163	
XX	XX	Percent Similarity:	73.94%	Conservative:	61	

Best Local Similarity:	55.45%	Mismatches:	84
Query Match:	40.23%	Indels:	2
DB:	22	Gaps:	1
US-09-761-640-4 (-1-471) x AAD12965 (-1-1949)			
QY 142	ASPIGLUTHTHValLeuLeuGlyValAspPheProAspSerSerSerProSerCysThLeu	161	
Db 32	GAGGAGAAATATCTTGCTGGAGTGGACTTTCAGTAAGAAAGTAAACCTGCACCATT	91	
QY 162	GlyLeuValLeuProLeuThrSerAspPheCysIleValLeuAspGlyAspGlyLeuPhe	181	
Db 92	GGGATGGTTCTCCGACTGTGGAGGACCGAAATCCACCTTGATGGAGATGGGTTC	151	
QY 182	SerValThrSerGlyGlyGlnSerArgIlePheIleProIleSerIleGlnThrMetTrp	201	
Db 132	AGCGTAGACACACGAGGAAGGATGCACATATTATTAAGCTGTGTCTGTCCAGGCCATGTG	211	
QY 202	AlaThrLeuGlnValLeuHisGlnAlaCysGlnAlaLeuGlySerGlyLeuValPro	221	
Db 212	TCTGCCCTCGAGGTCCTTCACAAAGGCGCCGCAAGTGGCCGAGGACCAACTACTTCCC	271	
QY 222	GlyGlySerAlaLeuThrThrAlaSerHisIleGlyGlnValArgLeuAsnSerGlnGlnSer	241	
Db 272	GGGGGTGAGCTCTCATCTGGAGCTACCTACATATAGAGCTGACATCACTCCGAGCAAGC	331	
QY 242	CysLeuAsnGlnIleTrpThrAlaMetAlaAspLeuGlnSerLeuArgPro-----ProSer	265	
Db 332	TGCATCAACGAGTGAACGCCATCGAGAGCTTGGAGTTCACGGCCGCCAGTCCCGCGC	391	
QY 260	AlaGluProGlyGlySerSerGlnGlnGlnMetGlnAlaIleArgAlaGluLeu	279	
Db 392	CTATTGTGGACAACCCCACTGGAAGGGGAAGGACCGAGCCGCTCATCAAAAGCCAACTC	451	
QY 280	TrpIleValLeuAspValSerAspLeuGlnSerValThrSerGlyGlnIleArgGlnAla	299	
Db 452	CGAACCATCATGATGAGCCAGGACATCTAAGAAATGTGATCTCCAAAGAGATTCGTAATGA	511	
QY 300	LeuGluLeuArgGlyGlyLeuProLeuGlnGlnIleGlyArgAspPheIleAspAsnGlnMet	319	
Db 512	TTAGGAACAACAGATGAATTGTAAGTGAAGGAACTCAAGAAATTTATAGACAATGAGATG	571	
QY 330	LeuLeuLeuValAlaGlnIleArgAspArgAlaSerArgIlePheProHisLeuIleGly	359	
Db 572	CTACTTATCTGGGACAGAGATGGACAAGCCCTCCCTTATCTTCGATCATCTTTATCTCGC	631	
QY 340	SerGluTrpAsnAlaAlaAsnLeuGlnIleGlnIleGlnArgAsnArgValThrHisIleLeu	358	
Db 632	TCTGATATGGATATGCTTCACAACTGTGAGGAACCTGCAGGGCTGAGGTGATTACATTTTA	691	
QY 360	AsnMetAlaArgGlnIleAspAsnPheIleProGluIleArgPheThrIleGlyHisAsnValArg	379	
Db 692	NATGTTACCAAGAAATGATATTTTTTTCTCTGGCTTATTTGCAATTCATACATCCGA	751	
QY 380	LeuTrpAspGlnGlnSerAlaGlnLeuLeuProHisTrpIleGlnThrHisArgPheIle	399	
Db 752	GCTCAGATGGAAGAACACAGACCTCTCCGCCACAGGAATGAAGGGTATCATTTATA	811	
QY 400	GlnAlaIleAlaArgAlaGlnGlyThrHisValLeuValHisCysIleSmetGlyValSerArg	419	
Db 812	AACAAAGGAAGAGAACCATTCACAGTGCCTGTGTGCATTGCCAAAAGGGCGTGAAGTCG	871	
QY 420	SerAlaAlaThrValLeuAlaIleValAlaMetCysGlnIleGlyCysSerLeuGlnAla	439	
Db 872	TGGGCTCCACAGTATAGCTTATGCAATTAAGGAATTTGGCTGGCTCTGGAAAAGACA	931	
QY 440	LeuArgHisGlnGlnLeuArgProIleAlaArgProAsnProGlyPheLeuArgGln	459	
Db 932	TATACTATATGTAACACGAACGCCGATCAACGGGCCCAACGGGGCTTTATAGAGCAG	991	
QY 460	LeuGlnIleIleGlyGlnGlyIleLeuThrAla 469		
Db 992	CTGTCTGAGTATGAAGCATCTTGGAATCA 1021		

RESULT 15
ABL10739
ID ABL10739 standard; cDNA; 4467 BP.
XX
AC ABL10739;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 26699.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KM pharmaceutical; gene; ss.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
XX
DR P-PSDB; ABB66636.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
PS Claim 1; SEQ ID NO 26699; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
XX sequences (AB101840-AB16175) and the encoded proteins
XX (ABB57737-ABB72072).
XX
CC The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 4467 BP; 1157 A; 1283 C; 1219 G; 808 T; 0 other;
SQ
Alignment Scores:
Pred. No.: 3.75e-67 Length: 4467
Score: 918.50 Matches: 205
Percent Similarity: 53.23% Conservative: 83
Best Local Similarity: 37.89% Mismatches: 164
Query Match: 37.71% Indels: 89
DB: 23 Gaps: 7
US-09-761-640-4 (1-471) x ABL10739 (1-4467)
QY 1 MetalaleValThrValSerArgSerProGlySerGlyAla---SerThrProVal 19
DB 460 ATGGCGCTGGTACGAGCGCTCTCCCACTGCGGCGCTCCGCTCGAATTCGGAT 519
QY 20 GlyProTTPASPGLNALAValAlaInArgArgSerArgLeuGlnArgArgInSerPheAla 39
DB 520 GCGCAGTCCGAGACGACGACGACATACCAAGAAAGCAGTCGACGACGACGCTCTT 579
QY 40 ValLeuArgGlyAlaValLeuGlnLeuGlnAspGlyGlyAspAsnAspAlaAlaGlu 59
DB 580 GCGCGCAAGGAAACCGCTGTACTG----- 606

QY 60 AlaSerSerGluProThrGluAlaProSerGlnGlnGlnLeuHisGlyAspGlnThr 79
DB 607 GCTCTTAAGCATATCCSCCCTGACGCAAAAGCAGAGCTCCCTGATGACGACGAC 666
QY 80 AspPheGlyGlnGlySerGlnSerProGlnGlnGlnGlnGlnGlnGlnGlnHis 99
DB 667 GCGTCACATAACAGTACCAAAAGC-----AACAACTGGACATCCACCTCCGCA 720
QY 100 LeuMetValGlnLeuLeuArgProGlnAspPheAlaValAlaGlnLeuGlnAla 119
DB 721 TCAATGTTTAACTGTTTCAGCGGACGACGCTAAAGATGAGCTGTCAAGTGAATCA 780
QY 120 ProArgProProArgLeuArgTyrLeuLeuValValSer----- 132
DB 781 CAGCGGTCCAAACCGCAGCGGATGCTGATGATGATGATGATGATGATGATGATGAT 840
QY 132 ----- 132
DB 841 ACCAGCAGCGTCCGCGCCACCGGATCATGCTCACACATCGGTAAAGTTGTTGATCC 900
QY 132 ----- 132
DB 901 GCGGGAACAAGTGTGTCACCTCCCGCGCGTCCCGCAGCAGCTGCTGCTGATG 960
QY 133 -----Thr 133
DB 961 CAGACCGCCACAGAGCGCCAGCAATGTGTATAACAGCAGCAAGAAAGCAACAG 1020
QY 134 ArgGlnGlyGlnGlyLeuSer-----GlnAspGlnThrValLeuLeuGlyAla 150
DB 1021 GCGCGCGCGGACACAAAGATACCTCGGGCAGAGAGATCTCGCTGCGCTAAG 1080
QY 151 PheProAspSerSerSerProSerCysThrLeuGlnGlyLeuValLeuProLeuPhe 170
DB 1081 TCCACAGCAGCGGAC-----ACATTGGCTGTGTGTGATCCATCGCGGAC 1128
QY 171 ThrGlnValTyrLeuAspGlyAspGlyGlyPheSerValThrSerGlyGlnSerArg 190
DB 1129 ACCACAATTCACCTGATGCGGATGCGGATGATGATGATGATGATGATGATGAT 1188
QY 191 IlePheIleProIleSerIleGlnThrMetTrpAlaThrLeuGlnValLeuHisGln 210
DB 1189 AATTTTAAACCGCTGCTCATGTCAGCCATGTCGCTGCGCTGACACGCTGACAAAGTG 1248
QY 211 CysGlnAlaAlaLeuGlnGlySerGlyLeuValProGlyGlySerAlaLeuThrPhe 230
DB 1249 TCGAATAAAGCTCTGAGAGAACACTTATGCGACGTGACCTCCACGACGCTATCC 1308
QY 231 HisTyrGlnIleArgLeuAsnSerGlnGlnSerCysLeuAsnGlnTrpThrAlaMet 250
DB 1309 ACCTACGAGAGCGCATGATGATGATGATGATGATGATGATGATGATGATGATG 1368
QY 251 AspLeuGlnSerLeuArgProProSerAlaGluPro-----GlyGlySerSerGln 268
DB 1369 GCGCTCGAAGCGCGGACGACCTTCTGCTGATGCGCAACGAAACCCCGCAAAAG 1428
QY 269 GlnGlnMetGlnGlnAlaIleArgAlaGlnLeuTrpIleValLeuAspValSerPhe 288
DB 1429 GAGGAACCGAAAGTCTATTAGATGAAGCTGAAGCCATATGATGATGATGATG 1488
QY 289 GlnSerValThrSerIleGlnIleArgGlnAlaLeuGlnAlaLeuArgLeuGlnLeu 308
DB 1489 GACGAGGTACCTCCAAAGTATATACCGGACCTTCGAGAGATCTCGACATGATTTG 1548
QY 309 GlnGlnTyrArgAspPheIleAspAsnGlnMetLeuLeuLeuValAlaGlnArgPhe 328
DB 1549 GCGGAGTACAAATCGTTTTCACGAAATGCTGTTATTTCTAGGCCAAATGGAATGCC 1608
QY 329 AlaSerArgIlePheProHisLeuTrpLeuGlnGlySerGlnTrpAspAlaAlaAsnLeu 348
DB 1609 CCCACCAAGATTTTGAGACGACGTATCTGTGGATCAGATGGAATGCCGCAACTTGGAG 1668

THIS PAGE BLANK (uspto)

GenCoife version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model.

Run on: April 11, 2003, 04:02:23 ; Search time 159 Seconds
(without alignments)
908.458 Million cell updates/sec

Title: US-09-761-640-4
Perfect score: 2436
Sequence: 1 MALVTVSRSPSGSGASTPVG.....PNCFLRLQLQIYGLTLART 471

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 15338381 residues
Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL-frame+ p2n model -DEV-rlh
-Q-/cgn2_1/USFPO.spool/US09761640/runtac.08042003_141437_20394/app_query.fasta-1.647
-DB-issued.patents.NA -OFMT-fast1 -SUFFIX-rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS-bits -START=1 -END=1 -MATRIX-blosum62 -TRANS-human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE-pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE-LOCAL -OUTFMT-pto -NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USRR-US09761640.ecgn_1.1.66_@runac.08042003_141437_20394 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAL -LARGEOUTR -NEG_SCORES=0 -WAIT -LONGLOG -DEV-TIMEOUT=120
-WARN_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: Issued Patents.NA.*
2: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*
3: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6A.COMB.seq.*
5: /cgn2_6/ptodata/2/ina/6B.COMB.seq.*
6: /cgn2_6/ptodata/2/ina/PCPUS.COMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	219.5	9.0	1987	2	US-08-990-379-1 Sequence 1, Appli
2	216	8.9	531	4	US-09-704-139-3 Sequence 3, Appli
3	216	8.9	1390	4	US-09-704-139-1 Sequence 1, Appli
4	216	8.9	1993	2	US-08-990-379-2 Sequence 2, Appli
5	211	8.7	1691	3	US-09-013-881-12 Sequence 1, Appli
6	207	8.5	1238	2	US-08-530-290-11 Sequence 11, Appli
7	194	8.0	912	4	US-09-163-833-1 Sequence 3, Appli
8	193	7.9	594	4	US-09-163-833-3 Sequence 6, Appli
9	187	7.7	789	4	US-09-164-193-6 Sequence 6, Appli
10	187	7.7	789	4	US-09-221-448A-6 Sequence 6, Appli
11	187	7.7	1016	4	US-09-164-193-4 Sequence 4, Appli
12	187	7.7	1016	4	US-09-221-448A-4 Sequence 4, Appli

13	186.5	7.7	861	1	US-07-988-273-1 Sequence 1, Appli
14	186.5	7.7	861	5	PCR-US93-12019-1 Sequence 1, Appli
15	185.5	7.6	1358	3	US-09-013-881-11 Sequence 11, Appli
16	182	7.5	1729	4	US-09-045-973-6 Sequence 6, Appli
17	182	7.5	1729	4	US-09-045-973-6 Sequence 9, Appli
18	136	5.6	821	2	US-08-530-290-9 Sequence 70, Appli
19	134	5.5	821	3	US-08-933-750C-70 Sequence 70, Appli
20	123.5	5.1	7295	2	US-09-234-613-70 Sequence 15, Appli
21	113.5	4.7	1098	3	US-08-487-826B-15 Sequence 5, Appli
22	113.5	4.7	1098	4	US-09-497-897-5 Sequence 2, Appli
23	110.5	4.5	71989	4	US-09-443-501A-2 Sequence 2, Appli
24	109.5	4.5	35828	4	US-09-449-218D-17 Sequence 17, Appli
25	108.5	4.5	80161	3	US-09-036-987A-1 Sequence 1, Appli
26	108.5	4.5	80161	4	US-09-370-700-1 Sequence 1, Appli
27	108	4.4	4031	1	US-08-159-784-1 Sequence 1, Appli
28	107	4.4	237	2	US-08-530-290-5 Sequence 5, Appli
29	107	4.4	5494	4	US-08-910-864-10 Sequence 10, Appli
30	107	4.4	5510	1	US-08-123-161A-7 Sequence 7, Appli
31	107	4.4	5510	4	US-08-483-278-7 Sequence 1, Appli
32	106.5	4.4	15611	4	US-09-441-340-1 Sequence 1, Appli
33	106.5	4.4	68750	3	US-09-335-409-1 Sequence 1, Appli
34	106.5	4.4	68750	4	US-09-568-102-1 Sequence 1, Appli
35	106.5	4.4	68750	4	US-09-567-969-1 Sequence 1, Appli
36	106.5	4.4	68750	4	US-09-568-480-1 Sequence 1, Appli
37	106.5	4.4	68750	4	US-09-568-480-1 Sequence 1, Appli
38	106.5	4.4	68750	4	US-09-568-472-1 Sequence 1, Appli
39	106.5	4.4	68750	4	US-09-567-899-1 Sequence 1, Appli
40	106	4.4	4456	4	US-09-095-443-1 Sequence 1, Appli
41	105	4.3	50341	1	US-08-247-901C-1 Sequence 1, Appli
42	105	4.3	50341	2	US-09-075-904-1 Sequence 1, Appli
43	105	4.3	52297	4	US-09-426-436-1 Sequence 1, Appli
44	105	4.3	52297	4	US-08-705-557-1 Sequence 1, Appli
45	103	4.2	3288	2	US-08-620-694A-1 Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-990-379-1
Sequence 1, Application US/08990379
Patent No. 5998188
GENERAL INFORMATION:
APPLICANT: Stock, Phillip J
TITLE OF INVENTION: Mitogen Activated Protein Kinase Phosphatase cDNAs and
FILE REFERENCE: 4104-00032205A
CURRENT APPLICATION NUMBER: US/08/990.379
CURRENT FILING DATE: 1997-12-15
EARLIER APPLICATION NUMBER: PCT/US96/10402
EARLIER FILING DATE: 1996-06-14
EARLIER APPLICATION NUMBER: 60/000.263
EARLIER FILING DATE: 1995-06-16
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 1987
TYPE: DNA
ORGANISM: Rattus norvegicus
US-08-990-379-1

Alignment Scores:

Pred. No.: 1.33e-12 Length: 1987
Score: 219.50 Matches: 112
Percent Similarity: 36.71% Conservatism: 51
Best Local Similarity: 25.23% Mismatches: 145
Query Match: 9.01% Indels: 137
DB: 2 Gaps: 17

US-09-761-640-4 (1-471) x US-08-990-379-1 (1-1987)

QY 50 AspGlyAspAsnAspAlaAlaGluAlaSerSerGluProThrGluLysAlaPro 69
|||||
:::
:::
|||


```

Db 581 GCGCTG-----586
Qy 245 GIUTRPTThralametalaspleuuserleuargproseralagluiproglygly 264
Db 587 CATATACATCGTCGCGGGGGGGCCAGGGGCTCGTAGCTTGACACAGAT-----637
Qy 265 serseglu-glnglglmetgluinalaleargalagluileutrypsvalleuas 284
Db 638 TCTGCCCGCGAGAGAGAGTGGCCCGCCCTCGGCTCTGAGCTTACTCGGCTGCAT 637
Qy 284 pValseraspLeuGluservalthrserlysgluilearglinalaleuGluleuargle 304
Db 698 CGTCCTACGATGAGCCGAGCCCGCCGAGACTCTCCGGAG-----GACAGCACAGT 751
Qy 304 uGluLeuProLeuGlnGlnltyrargaspheileasp---AsnGlnMetleuLeuVa 323
Db 752 GTCCGTGTCGTGTCAGAGCGTTGGCCGGAACGGAGCCGACAGACATCTGCTGCTTAA 811
Qy 323 lAlaglnArgaspArgalaser-----330
Db 812 AGGTGCTATGAGAGGTTTCTTCTGTAGTACCCAGATTCTGCTTAAACTAAGGCCCT 871
Qy 330 -----330
Db 872 GGGCGCCATCCACCCCGCTACCTCCAGACAAATGAGTCTGATCTGGGCTGCAG 931
Qy 331 -----ArgllepheProhileuty 337
Db 932 CTCCTGTGGAGCCCACTGCAGACACAGGGGGTCCCTGTGAGATCTTCTTCTCTTA 991
Qy 337 rLeuGlySerGluTrpAsnAlaAlaasnleuGluGluGlnArgAsnArgValThrH 357
Db 992 CCGCGGAGAGTCTTACACGCTGCGCGAGGAGCATGCTTATGCTTGGGAGTACAGGC 1051
Qy 357 sIleLeuAsnMetAlaArgGluLeuAspAsnPhetyrProGluArgPheThrTyrrHis 377
Db 1052 TCATCTGATGTCCTCTCAGACTGCCCAATCACTTGAGGAGCATTAACGACAAAGTG 1111
Qy 377 nValArgLeuTrpAspGluLeuSerAlaGlnLeuLeuProHistrlysgluThrHis 397
Db 1112 CACCCGGTAGAGATACCAACAGGCTGACAGTCTCTGTTTCAATGAGAACCATCGA 1171
Qy 397 gPheIleGluAlaAlaArgAlaGlnltyrHisValLeuValHisCysLysMetGly 417
Db 1172 ATCATATGACGCGAGTGAAGGACTGCCAGGGGAGTGTGTTACATGCCAGCCGCGCAT 1231
Qy 417 lSerArgSerAlaAlaThrValLeuAlaTyrrAlaMetLysGlnTyrglucysSerleu 437
Db 1232 CTCTAGATCAGCCACATCTGCTGGCTACGTGATGATGAAGAAACGGGTGAGGCTGGA 1291
Qy 437 uGlnAlaLeuArgHisValGlnGluLeuArgProIleAlaArgProAsnProGlyPhe 457
Db 1292 GAGAGGCTTTCGATGTTGCTCAAGCAGCGCGGTACATCATCTGCCCAACTTCAGCTCAT 1351
Qy 457 uArgGln---LeuGlnlIeTyrglnGlyleuThr 468
Db 1352 GGGCCAGTGTCTCAGTTCCAGTCTCAGTGTCTCAGC 1388

```

```

? STREET: 3174 Porter Drive
? CITY: Palo Alto
? STATE: CA
? COUNTRY: USA
? ZIP: 94304
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette
? COMPUTER: IBM Compatible
? OPERATING SYSTEM: DOS
? SOFTWARE: PASTESQ for Windows Version 2.0
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/013,881
? FILING DATE: HEREWITH
? CLASSIFICATION:
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER:
? FILING DATE:
? ATTORNEY/AGENT INFORMATION:
? NAME: BILLINGS, LUCY J.
? REGISTRATION NUMBER: 36,749
? REFERENCE/DOCKET NUMBER: PF-0470 US
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 650-855-0555
? TELEFAX: 650-845-4166
? TELEX:
? INFORMATION FOR SEQ ID NO: 12:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1691 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? IMMEDIATE SOURCE:
? LIBRARY: MUSCNOT02
? CLONE: 971204
?
US-09-013-881-12
?
Alignment Scores:
Pred. No.: 7,56e-12 Length: 1691
Score: 211.00 Matches: 56
Percent Similarity: 56.46% Conservative: 27
Best Local Similarity: 38.10% Mismatches: 56
Query Match: 8.66% Indels: 8
DB: 3 Gaps: 4
?
US-09-761-640-4 (1-471) x US-09-013-881-12 (1-1691)
Qy 327 ASPArgAlaserArgllepheProHistrlysgluLeuGlySerGluTrpAsnAlaAsn 346
Db 589 AACCATGCGCAGAGGTCGCGGCTGAGGAGCTTATCTGAGAGACAGACATGCTAACAA 648
Qy 347 LeuGluGluLeuGlnArgAsnArgValThrHisIleLeuAsnMetAlaArgGluIleAsp 366
Db 649 CGCGGGAGCTTGGCGCGCTGGGCAATCAGCAGCTCTCATATGCTGAGAGAGAGAGTGA 708
Qy 367 AsnPhetyrProGluArgPheThrTyrrHisAsnValArg---LeuTrpAspGluLeuSer 385
Db 709 CGAGGACAGCGCCAGAGGCTGATGAGGGGCTGGGCAATCCGTAACCTGGGTGAGCCAGC 768
Qy 386 AlaGlnLeuLeuProHistrlysgluThrHisArgPheIleGlnAlaAlaArgAlaGln 405
Db 769 TTTGACATGAGATCCATCTCCAGAGGCTGCCGACTTATCCACCGGGGCGGTGAGCCAG 828
Qy 406 ---GlyThrHisValLeuValHisCysLysMetGlyValSerArgSerAlaAlaThrVal 424
Db 829 CAGAGAGGAGAAATCCGTGGTGGTGTGGTGGGCGGTGAGAGACCGGACCGCTGGTA 888
Qy 425 LeuAlaTyrrAlaMetLysGlnTyrglucysSerLeuGluGlnAlaLeuArgHisValGln 444
Db 889 CTGGCTACCTCATGCTGTACACACACCTTACCTGCTGAGGAGCCATCAAGAAAGTCAAA 948
Qy 445 GluLeuArgProIleAlaArgProAsnProGlyPheLeuArgGlnleu-----460
Db 949 GAGCACCGAGGATCATC---CCCAACGGGGCTTCTGAGGAGGCTCTGGCCCTTGAGC 1005

```

QY 461 ---GlnIleTyrGlnGlyIle 466
Db 1006 CGCAGCGTGGCGCAGGCTGTG 1026

RESULT 6

US-08-530-290-11

Sequence 11, Application US/08530290
Patent No. 5958721

GENERAL INFORMATION:

APPLICANT: Marshall, Christopher John
APPLICANT: Ashworth, Alan
APPLICANT: Hughes, David Anthony
TITLE OF INVENTION: Methods for Screening of Substances for
TITLE OF INVENTION: Therapeutic Activity and Yeast for Use Therein
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/530.290
CLASSIFICATION: 435
FILING DATE: 14-DEC-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: WO PCT/GB94/00694
FILING DATE: 31-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9402573.1
FILING DATE: 10-FEB-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9307250.2
FILING DATE: 07-APR-1993

ATTORNEY/AGENT INFORMATION:

NAME: Bastian, Kevin L.
REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 084611-000000US
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:

LENGTH: 1238 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-530-290-11

Alignment Scores:

Pred. No.: 1.18e-11 Length: 1238
Score: 207.00 Matches: 105
Percent Similarity: 37.44% Conservative: 56
Best Local Similarity: 24.42% Mismatches: 164
Query Match: 8.50% Indels: 105
DB: 2 Gaps: 17

US-09-761-640-4 (1-471) x US-08-530-290-11 (1-1238)

QY 62 SerGluProThrGluAlaProSerGluGluLeuHisGlyAspGlnThrSer--- 80
Db 23 GCGCGCCGAGCGGCTGGTCCCGGCGAC-----CATGGTACGATGAGAGAGCT 73
QY 81 pheGlyGlnGlySerGlnSerProGlnGlnGlnGlu----- 93

Db 74 GCGGAGATGAGTGCAGTGTGCTCAAAAGCTGATGAACCGGAGCAGATGCGGCGG 133
QY 94 -----GlnArgGlnHisLeuHisLeuMetValGlnLeuLeuArg-----ProGln 108
Db 134 CGCGGCGGCGAGCGGCGAGCGACGACCTGGGGCTGCGGAGCGGCGGCGAGTGGCT 193
QY 109 AspAspIleArgLeuValAlaGlnLeuGlnAlaProArgProProArgLeu-----Arg 126
Db 194 GGTGACGTCAACACCGTCTTCTGGCGACAGCGCGGCTACATCTAGATTCGTCACACT 253
QY 127 TyrLeu-LeuValValSerThrArgGlnGlyGlyLeuSerGlnAspGlnThrValle 146
Db 254 GCGCTGTAAACACATCTGCGCGCGGCGGCTAAAGGCG----- 290
QY 146 uLeuGlyValAspPheProAspSerSerSerProSerGlyThrLeuGlyLeuValLeuPr 166
Db 291 -----TCGTGAGCGCTGGAGCAGATCTGCG 316
QY 166 OleuTrpSerAspThrGlnValTyrLeuAspGlyAspGlyPheSerValThrSerG1 186
Db 317 CGCGGAGGAGAGGTACGCGCGCGCTG---CGTCCGGCTCTACTGCGGCTATGCT 373
QY 186 yGlyGlnSerArgIlePheLysProIleSerIleGln-----ThMetTrpAlaTh 203
Db 374 CTACGACGAGCGAGCGCGCGCGCGAGCCTCCGCGAGCAGACGACCGCTGCTGCT 433
QY 203 rLeuGlnValLeuHisGlnAlaGlyGlnAlaLeuGlySerGlyLeuValProGlyL 223
Db 434 GGTGAGGCGGCTGCGCGCGAGCGCGCGACCC---GACATCTGCTGCTCAAAAGCGCG 490
QY 223 ySerAlaLeuThrTrpAlaSerHisTyrGlnGlnArgLeuAsnSerGlu-----Glnse 241
Db 491 C-----TATGAGAGCTTTCTCGAGTACCAGCAATT 523
QY 241 rCysLeuAsnGluTrpThrAlaMetAlaAspLeuGlnSerLeuArgProProSerAla-- 260
Db 524 CTGT---TCTAAACCAAGCCCTGGCAGCATTCCACCCCGGTTCCCGCAGCGCAC 580
QY 261 -GluProGlyGlySerSerGlnGlnGlnMetGlnGlnAlaIleArgAlaGlnLeuTr 280
Db 581 AGAGGCC----- 587
QY 280 pLysValLeuAspValSerAspLeuGlnSerValThrSerGlyIleArgGlnAlaLe 300
Db 588 -----TTGAGACTG---GACTGCGAGCTCC----- 608
QY 300 uGlnLeuArgLeuGlyLeuProLeuGlnGlnTyrArgAspPheIleAspAsnGlnMetLe 320
Db 609 -----TGTGGAGCCCATACAGCAGCAGAGGGTCTGTGAG----- 647
QY 320 uLeuLeuValAlaGlnArgAspArgAlaSerArgIlePheProHisLeuTyrLeuGlyse 340
Db 648 -----ATCTTCCCTCTCTACTCTGCGCAG 673
QY 340 rGluTrpAsnAlaAlaAsnLeuGlnGlnLeuGlnArgAsnArgValThrHisIleLeuAs 360
Db 674 TGCCTACATGCTGCGCGGAGAGACATGCTGAGAGCGCTGGGCGATCAGCGCTGTGTA 733
QY 360 mMetAlaArgGlnIleAspAsnPheTyrProGlnArgPheThrTyrHisAsnValArgLe 380
Db 734 TGTCTCTCCGAGCTGCGGAGACCATTTGAAGACACTATGACATGATGATGATGATGAT 793
QY 380 uTrpAspGlnGlnSerAlaGlnLeuLeuProHisTyrGlyGlnArgPheIleLe 400
Db 794 GGAAGATTAACCAAGCGCGCATGCTGCTGCTGATGAGAGCATGAGATGAGATGAGTA 853
QY 400 uAlaAlaArgAlaGlnGlyThrHisValLeuValHisCysLysMetGlyValSerArgse 420
Db 854 TGCCTGGAAGAGACTGCGGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 913
QY 420 rAlaAlaThrValLeuAlaTyrAlaMetLysGlnTyrGlyCysSerLeuGlnAlaLe 440
Db 914 GGCACACATCTGCTGCTGCTGCTGCTGATGATGAAGAACGGGTGAGGCTGAGAGAGGCTTT 973


```
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(789)
US-09-164-193-4

Alignment Scores:
Pred. No.: 9,37e-10 Length: 1016
Score: 187.00 Matches: 47
Percent Similarity: 53.03% Conservative: 23
Best Local Similarity: 35.61% Mismatches: 56
Query Match: 7.68% Indels: 6
DB: 4 Gaps: 2

US-09-761-640-4 (1-471) x US-09-164-193-4 (1-1016)
QY 327 ASPARGALSERARGILEPHEPROHISLEUTYRLEUGLYSERGLUTRPSNALAALAASN 346
DB 49 AACCATGCCGACGAGGTCTGGCCAGGCGCTTATCTCGAGACGACGATGGCTAACAC 108
QY 347 LEUGLUGLLEUGLNAARGVATPHTHISILEUASMETGLAARGLUILEASP 366
DB 109 CGCCGGAGCTTCGCCCGCTGGGCATACGACGCTCTCAATGCTCACACAGCCGGTGG 168
QY 367 ASNPHTYRPROGLUARGPHE-----THYTHISASNYVALARGLEUTRP 381
DB 169 CGAGCGACCGCCGAGGCGCTATGAGGGCTGGGCATCCGCTACCTGGGTGTGAGGCCAC 228
QY 382 ASPLUGLUSERALAGLNEULEUPROHISTRPYLSGLUTHTHISARGPHEILEGUALA 401
DB 229 GACTGCCAGCGCTTGACATGAGCATCCATCCAGCGGCTGCCAGCTTCATCCACCGG 288
QY 402 ALAARGALAGLN--GLYTHRHISVALLEUVALHISCYLSYMETGLYALSERARGSER 420
DB 289 GCGCTGACGACGACGAGGAGGAAGATCCTGTCATCTGCTGTGGCGGTGAGCCGATCC 348
QY 421 ALAALATHVALLEUALATYRALAMETLYSGINTYRGLUCYSERLEUGLINALALEU 440
DB 349 GCCACCGCTGACTGCGCTACCTCATGCTGTACACACACCTTACCTCGTGAGGCCATC 408
QY 441 ARGHISVALGLNGLLEUARGPROILEALARGPRO 452
DB 409 AAGAAAGTCAAGACCAACGAGGAGGAGGCCGAGCC 444

RESULT 12
US-09-221-448A-4
; Sequence 4, Application US/09221448A
; Patent No. 6436685
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPTP PROTEIN MOLECULES AND USES THEREFOR (AS AMENDED)
; FILE REFERENCE: MN-051DV1
; CURRENT APPLICATION NUMBER: US/09/221,448A
; CURRENT FILING DATE: 1998-12-28
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1016
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(789)
US-09-221-448A-4

Alignment Scores:
Pred. No.: 9,37e-10 Length: 1016
Score: 187.00 Matches: 47
Percent Similarity: 53.03% Conservative: 23
Best Local Similarity: 35.61% Mismatches: 56
Query Match: 7.68% Indels: 6
DB: 4 Gaps: 2
```

```
US-09-761-640-4 (1-471) x US-09-221-448A-4 (1-1016)
QY 327 ASPARGALSERARGILEPHEPROHISLEUTYRLEUGLYSERGLUTRPSNALAALAASN 346
DB 49 AACCATGCCGACGAGGTCTGGCCAGGCGCTTATCTCGAGACGACGATGGCTAACAC 108
QY 347 LEUGLUGLLEUGLNAARGVATPHTHISILEUASMETGLAARGLUILEASP 366
DB 109 CGCCGGAGCTTCGCCCGCTGGGCATACGACGCTCTCAATGCTCACACAGCCGGTGG 168
QY 367 ASNPHTYRPROGLUARGPHE-----THYTHISASNYVALARGLEUTRP 381
DB 169 CGAGCGACCGCCGAGGCGCTATGAGGGCTGGGCATCCGCTACCTGGGTGTGAGGCCAC 228
QY 382 ASPLUGLUSERALAGLNEULEUPROHISTRPYLSGLUTHTHISARGPHEILEGUALA 401
DB 229 GACTGCCAGCGCTTGACATGAGCATCCATCCAGCGGCTGCCAGCTTCATCCACCGG 288
QY 402 ALAARGALAGLN--GLYTHRHISVALLEUVALHISCYLSYMETGLYALSERARGSER 420
DB 289 GCGCTGACGACGACGAGGAGGAAGATCCTGTCATCTGCTGTGGCGGTGAGCCGATCC 348
QY 421 ALAALATHVALLEUALATYRALAMETLYSGINTYRGLUCYSERLEUGLINALALEU 440
DB 349 GCCACCGCTGACTGCGCTACCTCATGCTGTACACACACCTTACCTCGTGAGGCCATC 408
QY 441 ARGHISVALGLNGLLEUARGPROILEALARGPRO 452
DB 409 AAGAAAGTCAAGACCAACGAGGAGGAGGCCGAGCC 444

RESULT 13
US-07-988-273-1
; Sequence 1, Application US/07988273
; Patent No. 5512434
; GENERAL INFORMATION:
; APPLICANT: ARMONSON, Stuart A.
; APPLICANT: BOTTARO, Donald P.
; APPLICANT: ISHIBASHI, Toshio
; APPLICANT: MIKI, Toru
; TITLE OF INVENTION: EXPRESSION CLONING OF A HUMAN
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/988,273
; FILING DATE: 19921214
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 40399/182 NIHD
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 861 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
```

```

; LOCATION: 29..586
US-07-988-273-1

Alignment Scores:
Pred. No.: 8.11e-10 Length: 861
Score: 186.50 Matches: 51
Percent Similarity: 49.09% Conservative: 30
Best Local Similarity: 30.91% Mismatches: 73
Query Match: 7.66% Indels: 11
DB: 1 Gaps: 4

US-09-761-640-4 (1-471) x US-07-988-273-1 (1-861)

QY 306 LeupProleuGlnGlnTyrArgAspPheilleaspasnglMetleuLeuValAlaGln 325
   ||| :||| ||| :||| :||| :|||
Db 47 CTCCTGGTGCAGAGATCTCAACGACCTGCTCTCGACGGCAGCGGCTGTACAGCCTCCG 106

QY 326 ArgAspArgAlaSerArgIlePheProHisLeuTyrLeuGlySerGluTrpAsnAlaIa 345
   :||| :||| :||| :||| :||| :|||
Db 107 AGCCAGCCCTGCACAGAGTCAACCCCGGATCTACGTGGCAGACGCTGTGGCTCAG 166

QY 346 AsnLeuGlnGluLeuGlnArgAsnArgValThrHisIleLeuAsnMetAlaArgGluIle 365
   :||| :||| :||| :||| :||| :|||
Db 167 GACATCCCAAGCTGCAGAACTAGACGACCATGTGCTGAAGCGGCTGAGGGCAGG 226

QY 366 Asp-----AsnPheTyrProGluArg---PheThrTyrHis 376
   :||| :||| :||| :||| :||| :|||
Db 227 TCCTTCATGACAGTCAACACCAATGCCAATCTTCTACAGAGACCTCGGATCATACACTG 286

QY 377 AsnValArgLeuTrpAspGlnGluSerAlaGlnLeuLeuProHisTrpIleGluThrHis 396
   :||| :||| :||| :||| :||| :|||
Db 287 GGCATCAAGGCCAAGCAGACACAGAGTTCACACCTGACCGCTTCTGAAGGGCTGCC 346

QY 397 ArgPheIleGluAlaAlaArgAlaGln---GlyThrHisValLeuValHisCysIleMet 415
   ||| :||| :||| :||| :||| :|||
Db 347 GACCTCATTTGACACAGGCTTTGGCTCAAAAGAAATGGCCGGGCTGTCTGCACCTGCCGGGAA 406

QY 416 GlyValSerArgSerAlaAlaThrValLeuAlaTyrAlaMetLeuGlnTyrGluCysSer 435
   ||| :||| :||| :||| :||| :|||
Db 407 GGTATTACCCGCTGCCCAAGCTAGTTATTCGCTTACCTCATGATCGGACAGAAAGATGGAC 466

QY 436 LeuGlnGlnAlaLeuArgHisValGlnGlnLeuArgProIleAlaArgProAsnProGly 455
   :||| :||| :||| :||| :||| :|||
Db 467 GTCAAGTCTGCCCTGAGCATGCTGAGCGCAGAACCTGAGATCGGC---CCCAACGATGGC 523

QY 456 PheLeuArgGlnLeu 460
   ||| :||| :||| :||| :||| :|||
Db 524 TTCCTGGCCCAAGCTC 538

RESULT 14
PCT-US93-12019-1
; Sequence 1, Application PC/TUS9312019
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: EXPRESSION CLONING OF A HUMAN
; TITLE OF INVENTION: PHOSPHATASE
; NUMBER OF SEQUENCES: 7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/12019
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/988,273
; FILING DATE: 14-DEC-1992
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 861 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double

```

```

; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 29..586
PCT-US93-12019-1

Alignment Scores:
Pred. No.: 8.11e-10 Length: 861
Score: 186.50 Matches: 51
Percent Similarity: 49.09% Conservative: 30
Best Local Similarity: 30.91% Mismatches: 73
Query Match: 7.66% Indels: 11
DB: 5 Gaps: 4

US-09-761-640-4 (1-471) x PCT-US93-12019-1 (1-861)

QY 306 LeupProleuGlnGlnTyrArgAspPheilleaspasnglMetleuLeuValAlaGln 325
   ||| :||| ||| :||| :||| :|||
Db 47 CTCCTGGTGCAGAGATCTCAACGACCTGCTCTCGACGGCAGCGGCTGTACAGCCTCCG 106

QY 326 ArgAspArgAlaSerArgIlePheProHisLeuTyrLeuGlySerGluTrpAsnAlaIa 345
   :||| :||| :||| :||| :||| :|||
Db 107 AGCCAGCCCTGCACAGAGTCAACCCCGGATCTACGTGGCAGACGCTGTGGCTCAG 166

QY 346 AsnLeuGlnGluLeuGlnArgAsnArgValThrHisIleLeuAsnMetAlaArgGluIle 365
   :||| :||| :||| :||| :||| :|||
Db 167 GACATCCCAAGCTGCAGAACTAGACGACCATGTGCTGAAGCGGCTGAGGGCAGG 226

QY 366 Asp-----AsnPheTyrProGluArg---PheThrTyrHis 376
   :||| :||| :||| :||| :||| :|||
Db 227 TCCTTCATGACAGTCAACACCAATGCCAATCTTCTACAGAGACCTCGGATCATACACTG 286

QY 377 AsnValArgLeuTrpAspGlnGluSerAlaGlnLeuLeuProHisTrpIleGluThrHis 396
   :||| :||| :||| :||| :||| :|||
Db 287 GGCATCAAGGCCAAGCAGACACAGAGTTCACACCTGACCGCTTCTGAAGGGCTGCC 346

QY 397 ArgPheIleGluAlaAlaArgAlaGln---GlyThrHisValLeuValHisCysIleMet 415
   ||| :||| :||| :||| :||| :|||
Db 347 GACCTCATTTGACACAGGCTTTGGCTCAAAAGAAATGGCCGGGCTGTCTGCACCTGCCGGGAA 406

QY 416 GlyValSerArgSerAlaAlaThrValLeuAlaTyrAlaMetLeuGlnTyrGluCysSer 435
   ||| :||| :||| :||| :||| :|||
Db 407 GGTATTACCCGCTGCCCAAGCTAGTTATTCGCTTACCTCATGATCGGACAGAAAGATGGAC 466

QY 436 LeuGlnGlnAlaLeuArgHisValGlnGlnLeuArgProIleAlaArgProAsnProGly 455
   :||| :||| :||| :||| :||| :|||
Db 467 GTCAAGTCTGCCCTGAGCATGCTGAGCGCAGAACCTGAGATCGGC---CCCAACGATGGC 523

QY 456 PheLeuArgGlnLeu 460
   ||| :||| :||| :||| :||| :|||
Db 524 TTCCTGGCCCAAGCTC 538

RESULT 15
US-09-013-881-11
; Sequence 11, Application US/09013881
; Patent No. 6132964
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: HUMAN HYDROLASE-LIKE MOLECULES
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESS: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:

```

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/013.881
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: BILLINGS, LUCY J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0470 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1358 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: TMLR3D102
CLONE: 507537
US-09-013-881-11

Alignment Scores:
Pred. No.: 2.1e-09 Length: 1358
Score: 185.50 Matches: 63
Percent Similarity: 43.18% Conservative: 32
Best Local Similarity: 28.64% Mismatches: 98
Query Match: 7.61% Indels: 27
DB: 3 Gaps: 4

US-09-761-640-4 (1-471) x US-09-013-881-11 (1-1358)

QY 256 ArgProProSerAlaGluProGlyGlySerSerGluGlnGlnMetGluGlnAlaIle 275
DB 9 CGTCCGAGGCCGCGCGCTGTCTGTGCGCGCCATGTGGAGGCGAT-----GTT 53
QY 276 ArgAlaGluLeuTrpLysValLeuAspValSerAspLeuGluSerValThrSerLysGlu 295
DB 54 CTAGCCGGAGTCTACTCGATG----- 74
QY 296 IleArgGlnAlaLeuGluLeuArgLeuGlyLeuProLeuGlnGlnIleTyrArgAspPhe 315
DB 75 GTAGGCGCAGAGCCGCTGTCTGTGCGCGCCATGTGGAGGCTCCGGCCGAGT 134
QY 316 Asp-----AsnGlnMetLeuLeuValAlaGlnArgAspArgAlaSer 330
DB 135 GATGGCTGCGACCTGACCAACCCAGCCAGCAGACAGTCAGCTGCCGGCGATGCTG 194
QY 331 ArgIlePheProHisLeuTyrLeuGlySerGluTyrPasnAlaAlaAsnLeuGluGlu 350
DB 195 GAAGTGCACACGAGATTGATTTGCGTGGGCGCGCGCGTGGCGAGCCAGATCAGCTG 254
QY 351 GlnArgAsnArgValThrHisIleLeuAsnMetAlaArgGluIleAspAsnPheTyrPro 370
DB 255 ACGGAAGCGCGGATCAGCGCCGCTGCTAACAGTGCAGTCCGAGAGCCAGCTTCAAGCGC 314
QY 371 GluArgPheThrTyrHisAsnValArgLeuTyr-----AspGluGluSerAla 386
DB 315 GGGCGTGGGCTGAGGAGTATAGCGGCTTCTGTCGACGCTGACCGCTGACCAACCCGAGACG 374
QY 387 GlnLeuLeuProHisTyrLysGluThrHisArgPheIleGluAlaAlaArgAlaGlnGly 406
DB 375 GACCTACTCAGCATCGACCGGTCGTCGGCTTTCATCGGTACGCGCGCGTGAAGGC 434
QY 407 ThrHisValLeuValHisCysLysMetGlyValSerArgSerAlaAlaThrValLeuAla 426

DB 435 CGTGGGTGTGTGGTGCACGTCTCATGCGAGAGTCAGTCAAGTGTGGCCATTAATTAATCTGCT 494
QY 427 TyrAlaMetLysGlnTyrGluCysSerLeuGlnGlnAlaLeuArgHisValGlnGluLeu 446
DB 495 TTTCATGTAAGACGTGACCACTTCCCTTTGAAGAAAGCCTATGAAAGGCTCCAGATTCTC 554
QY 447 ArgProIleAlaArgProAsnProGlyPheLeuArgGlnLeuGlnIleTyrGlnGlyLe 466
DB 555 AAACGAGGCTTACATGATGAGGGGTTTGAGTGGCAACTGAATTTATACGAGGCAATG 614

Search completed: April 12, 2003, 10:55:32
Job time : 176 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 11, 2003, 03:44:27 ; Search time 3860 Seconds

(without alignments)
1976.185 Million cell updates/sec

Title: US-09-761-640-4
Perfect score: 2436
Sequence: 1 MALVTVSRSPGSGASTPVG.....PNEGFLRLQIYQILHART 471

Scoring table:
BIOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n.model -DEV=x1h
-O=/cgn2_1/USPFO.spool/US09761640/rnact_08042003_141435_20370/app_query.fasta.1.647
-DB=EST -QPM=fastcap -SUFFIX=1st -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=BIOSUM62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=us09761640.qcgn.1.1.3256.6runat.08042003_141435_20370 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAB -LARGEOBURY -NES_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN_TIMEOUT=30 -THREHOLD=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6 -FCGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estlun:*
5: em_estlov:*
6: em_estlpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estlun:*
16: em_estlun:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrl:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1413	58.0	1004	14	BM910493
2	1350	55.4	991	13	BM471207
3	1325	54.4	875	12	BG752969
4	1318	54.1	1024	13	BM560192
5	1301.5	53.4	812	12	BG750408
6	1202	49.3	790	12	BG326709
7	1147.5	47.1	846	10	BE563259
8	1147.5	47.1	910	12	BG169297
9	1144.5	47.0	726	12	BE869222
10	1091.5	44.8	919	13	BE737087
11	1089.5	44.7	943	13	BM563401
12	1085	44.5	932	12	BG340220
13	1042.5	42.8	897	12	BE873337
14	1009	41.4	663	12	BG753072
15	989.5	40.6	716	12	BE907514
16	982	40.3	803	13	B1105472
17	968	39.7	589	9	AL602806
18	966	39.7	791	12	BE881145
19	959.5	39.4	892	12	BF971565
20	916.5	37.6	1228	14	BM921421
21	910	37.4	969	10	BE563116
22	900.5	37.0	704	14	BQ108910
23	899	36.9	846	13	B1820326
24	897	36.8	937	14	BQ718968
25	884.5	36.3	658	12	BG822776
26	870	35.7	863	12	BG749256
27	805	33.0	1120	12	B1555292
28	797.5	32.7	860	12	BE127735
29	770	31.6	652	10	BB660625
30	767	31.5	612	13	BM017088
31	767	31.5	1121	13	BM470627
32	755.5	31.0	761	12	BF579815
33	728	29.9	471	13	BM482726
34	702	28.8	424	10	AW326161
35	691	28.4	1125	14	BM809903
36	689	28.3	599	10	BE280203
37	689	28.3	1069	13	BM560223
38	688	28.2	743	13	BM017302
39	687	28.2	409	10	AW732634
40	684.5	28.1	771	13	B1146038
41	684	28.1	1158	14	BM910866
42	672	27.6	799	13	BG965294
43	662	27.2	629	14	BM785529
44	641	26.3	890	12	BE901875
45	631.5	25.9	986	14	BQ715279

ALIGNMENTS

RESULT 1
LOCUS BM910493
DEFINITION AGENCOURT.6703004 NIH_MGC_98 Homo sapiens cdna clone IMAGE:5455997
5' mRNA sequence.
ACCESSION BM910493
VERSION BM910493.1 GI:19360872
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1004)
NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE

JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC

CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://www.lnl.org>

Plate: LCM1948 row: g column: 22
High quality sequence stop: 725.

FEATURES

Source

Location/Qualifiers
1. .1004

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5453997"
/clone_lib="NIH-MGC_98"
/tissue_type="astrocytoma grade IV, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: brain; Vector: pOTB1; Site:1: XhoI; Site:2
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAC(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-CDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH-MGC Library."

```

BASE COUNT	211 a	299 c	324 g	167 t	3 others
ORIGIN					

Alignment Scores:

Pred. No.:	2,546,137	length:	1004
Score:	1413.00	Matches:	299
Percent Similarity:	77.69%	Conservative:	4
Best Local Similarity:	76.67%	Mismatches:	18
Query Match:	58,008	Indels:	70
DB:	14	Gaps:	4

US-09-761-640-4 (1-471) x BM910493 (1-1004)

QY	25	AlaValGlnAArgSerArgLeuGlnIrrArgGlnSerPheAlaValLeuArgGlyIa	44
Db	2	GGGGTCCAGGCCAAGAGAGCTGCACCTCCAGGAAGCCAGACCTTTGGCGTCTCCGGGGCT	61
QY	45	ValLeuGlyLeuGlnAspGlyGlyAspAsnAspAspAlaAlaGluAlaSerSerGluPro	64
Db	62	GTCCTGGAGCTGCAGAGATGAGAGGGGACCAATGATGATGCAGACAGAGCCATGTTGAGCCA	122
QY	65	ThrGluLysAlaProSerGluGluGluLeuIleuHisGlyAspGlnThrAspPheGlyGlnGly	84
Db	122	ACAAGAAGAGGCCCGAGGTGAGAGAGAGCTCCACGGGGCCACAGACAGATTTGGGCAMGA	182
QY	85	SerGlnSerProGlnLysGlnGluGlnIrrGlnIrrGlnIrrHisLeuHisLeuMetValGlnLeu	104
Db	182	TCCAGAGTCCCCAGAAAGCAGAGAGAGAGAGCAGCAGCAGCCTGCAGCCTCATGTACGTACGTCG	242
QY	105	LeuArgProGlnAspAspIleArgLeuAlaAlaGlnLeuGlnAlaProArgProProArg	122
Db	242	CTGAGGGCCGCGAGATGCATCCGCTGTGACCCAGCTGGAGGCACCCCGGGCTCTCCCGG	302
QY	125	LeuArgTyrLeuLeuValValSerThrArgGluGlyGluGlyLeuSerGlnAspGluThr	144
Db	302	CTCCGCTACCTGCTGGTGAAGTTCTTACACGAAAGAGAGAGGTCTCGAGCCAGAGAGAGACG	362
QY	145	ValLeuLeuGlyValAlaAspPheProAspSerSerSerProSerCysThrLeuGlyLeuVal	164
Db	362	GTCCTCTGGGGCGTGATTTCCCTGCAGACAGAGCTCCCCACACTGCACACCTTGCGCTGGTC	422
QY	165	LeuProLeuIrrPsrAspThrGlnValTyrLeuAspGlyAspGlyGlyPheSerValThr	184
Db	422	TTTGGCCCTCTGGAGTGCACCCAGAGTGTACTTAATATGAGACGGGGGCTTCAGCGTACG	482

OY	185	SerGlyGLNserArgIlePheLeuProIleSerTIGLIThrMetThrAlaTrpIleu	204
Db	482	TCTGGGGGCMAAGCCGGATCTTCACAGGCCCATCTGCATGCACACCAATGTGG-----	532
OY	205	GlnValIleuHisGlnAlaCysGluAlaAlaIleuGlySerGlyLeuValProGlyGlySer	224
Db	532	-----	532
OY	225	AlaLeuThrTrpAlaSerHisTyrGlnGluArgLeuAsnSerGluInbertySleuAsn	244
Db	532	-----	532
OY	245	GluTrpTrpAlaMetAlaAspLeuGluSerLeuArgProProSerAlaGluProGlyGly	264
Db	532	-----	532
OY	265	SerSerGluGlnGlnGlnMetGluGlnAlaIleAlaGlyAlaGlyLeuTrpTyrValLeuAsp	284
Db	533	TCTTAGAACAAGACAGATGTCAGAGAGCGGATCCGTGTGACTGTGGAAAGTTTGAT	592
OY	285	ValSerAspLeuGluSerValThrSerTyrGluIleArgGlnAlaLeuGlnLeuArgLeu	304
Db	593	GTCACATGCACCTGGAGATGTCTACTTCCAAAGAAGATCCGCCAGGCTTGGACGTGCCCTTG	652
OY	305	GlyLeuProLeuGlnGlnTyrArgAspPheIleAspAsnGlnMetLeuLeuValAla	324
Db	653	GGGCTCCCCCTCCACAGTACCGTACTTCACTTCACGACMACCAATGCTGTGCTGGTGGCA	712
OY	325	GlnAlaGSPARGAlaSerArgIlePheProHisLeuTyrLeuGlySerGluTrpAsnAla	344
Db	713	CAGCGGACCGAGGCTCCCGCATCTTCCCACCCTCTACTGTGGCTCAGAGTGGAAACGA	772
OY	345	AlaAsnLeuGlnGlnLeuGlnAlaArgAsnArgValThrHisIleLeuAsnMetalArg-GI	364
Db	773	GCMAACCTGTGGAGACTGCAGAGGAACAAGGGTACCACCATCTTNGACATGTGGCGCGGGA	832
OY	364	IleAspAsnPheTyrProGluArgPheThrTyrHisAsnValArgLeuTrpAspGluGI	384
Db	833	GATTACACACTTCTACCTGAGCGCTTCACTACACACAG-GTGCGGCTTCTCGGGGATGA	891
OY	384	UserAla-----GlnLeuLeuProHisTrp--LysGluTrpHisArg-PheIleGluA	401
Db	892	AGGATCGGGCCAGACTGCTGCCCGCACCTGGAAAGAGAGACGACCGGCTTCATTGNAG	951
OY	401	Ia--AlaArgAlaGlnGlyThrHis	408
Db	952	GGCTTGCAGAGACACAAAGGCACAC	977
RESULT 2			
BM471207		991 bp	mRNA linear EST 05-FEB-2007
LOCUS			
DEFINITION	AGENCOURT_6478155 NIH_MGC_72 Homo sapiens cdna clone IMAGE:5563111		
ACCESSION			
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
TITLE	NIH-MGC http://mgs.ncl.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: gsp@rsfmail.nih.gov Tissue Procurement: ATCC/DCTP CDNA Library Preparation: Life Technologies, Inc. DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov		

Plate: LIAM1293 row: j column: 08
 High quality sequence stop: 592.
 Location/Qualifiers
 1. .991
 source

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="IMAGE:5563111"
 /clone_lib="NIH_MGC-72"
 /tissue_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: skin; Vector: pCMV-Sport6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 2 kb. Library constructed by Life
 Technologies."

BASE COUNT 199 a 303 c 329 g 159 t 1 others
 ORIGIN

Alignment Scores:

Pred. No.: 1e-130 Length: 991
 Score: 1350.00 Matches: 278
 Percent Similarity: 91.26% Conservative: 4
 Best Local Similarity: 89.97% Mismatches: 23
 Query Match: 55.42% Indels: 4
 DB: 13 Gaps: 2

US-09-761-640-4 (1-471) x BM471207 (1-991)

QY 1 MetAlaLeuValThrValSerArgSerProProGlySerGlyAlaSerThrProValGly 20
 Db 55 ATGGCCCTGGTGACAGAGAGCCGTTGCCCCGGGACGGCGCTCCACGCCGTGGGG 114
 QY 21 ProTrpAspGlnAlaValGlnArgSerArgLeuGlnArgGlnSerPheAlaVal 40
 Db 115 CCCGCGGACCGGGGGCTCCACGAGAGTGCATCCAGCAGGACAGCTTTCGGTG 174
 QY 41 LeuArgGlyAlaValLeuGlyLeuGlnAspGlyGlyAspAsnAspAspAlaAlaGlnAla 60
 Db 175 CTCGCGTGGGCTGCTCGGAGCTCAGAGATGAGAGGACATGATGATGACAGAGAGCC 234
 QY 61 SerSerGlnProThrGlnGlnAlaProSerGlnGlnGlnHisGlyAspGlnThrAsp 80
 Db 235 AGTTCTGAGCCCAAGAGAAAGCCCGGAGTGAAGAGGAGCTCCAGGGGACACAGAC 294
 QY 81 PheGlyGlnGlySerGlnSerProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 100
 Db 295 TTGCGGCGAGAGATCCACAGAGTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 354
 QY 101 MetValGlnLeuLeuArgProGlnAspAspIleArgLeuAlaAlaGlnLeuGlnAlaPro 120
 Db 355 ATGCTACAGCTGTAGAGCCGACAGATGACATCCGCTGGCAGCCAGCTGGAGGACCC 414
 QY 121 ArgProProArgLeuArgTyrLeuLeuValValSerThrArgGlnGlnGlnGlnGlnGln 140
 Db 415 CGGCGTCCCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 474
 QY 141 GlnAspGlnThrValLeuGlnGlyValAspPheProAspSerSerSerProSerCysThr 160
 Db 475 CAGAGTGAAGAGGCTCTCTGCGGTGATTTCCCTGACAGAGTCCCGCAGCTGCAC 534
 QY 161 LeuGlyLeuValLeuProLeuTyrSerAspThrGlnValTyrLeuAspGlyAspGly 180
 Db 535 CTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 594
 QY 181 PheSerValThrSerGlyGlnSerArgIlePheIysProIleSerIleGlnThrMet 200
 Db 595 TTCAGCGTGAAGTGTGGGCGAAAGCCGAGTCTTCAAGCCATCTCCAGAGCCATG 654
 QY 201 TrpAlaThrLeuGlnValLeuHisGlnAlaCysGlnAlaAlaLeuGlySerGlyLeu-Val 220
 Db 655 TGGGCGACACTCCAGGTTTTCACACAGCATGTGAGGAGAGCTTGAAGAGGGCTTGTAC 714
 QY 220 LProGlyGlySerAlaLeuThrTrpAlaSerHisTyrGlnGlnGlnGlnGlnGlnGlnGln 240

Db 715 CCCGGGTGGAGTGCCCTACCTGGGCGACCCACTACGAGAGAGACTTGAACTCCGACAA 774
 QY nSerCysLeuAsnGlnTrpThrAlaMetAlaAspLeuGlnSerLeuArgProProSerAl 260
 Db 775 GAGCTGCTCATATGAGGGAGCGGCTATGCGGACCTGGAGTCTTGCGGGCTCCGCCGCC 834
 QY 260 aGluProGlyGlySerSerGlnGlnGlnGlnMetGlnAlaAlaIleArgAlaGlnLeu- 279
 Db 835 CAACCTGGCGGGGCTCCGAAACAGAACGAAAGGAGCAGAGCGATCCCTGCTTAACCTG 894
 QY 280 --TrpValLeuAspValSerAspLeuGln--SerValThrSerIysGlnIleArg 298
 Db 895 TGGAAAAGGTTTGATGATTCAGGACCCGGAAGAGGTCCCTCCAAAAAACCCTCCG 954
 QY 298 InAlaLeuGlnLeuArgLeuGly 305
 Db 955 GGCTTCTGAACTTGCCCTGGGG 977

RESULT 3
 BG752969
 LOCUS 602732427F1 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:4876014 5',
 DEFINITION mRNA sequence.
 ACCESSION BG752969
 VERSION BG752969.1 GI:14063622
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 875)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://Image.lnl.gov

Plate: LCM1758 row: e column: 07
 High quality sequence stop: 818.
 Location/Qualifiers
 1. .875
 source

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="IMAGE:4876014"
 /clone_lib="NIH_MGC-43"
 /tissue_type="normal pigmented retinal epithelium"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: eye; Vector: pORF7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGCACGAG(G). Library constructed by Ling Hong
 in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library. 1"

BASE COUNT 184 a 267 c 275 g 149 t
 ORIGIN

Alignment Scores:
 Pred. No.: 3.41e-128 Length: 875
 Score: 1325.00 Matches: 276
 Percent Similarity: 94.86% Conservative: 12
 Best Local Similarity: 94.52% Mismatches: 12
 Query Match: 54.39% Indels: 6
 DB: 12 Gaps: 1

US-09-761-640-4 (1-471) x BG752969 (1-875)

```

QY 66 GltysAlaProSerGluGluLeuHisGlyAspGlnThrAspPheGlyGlnGlySer 85
   |||||
Db 1 GAGAAAGCCCCAGTAGAGAGAGCTCCAGGGAGACCAGACACTTCGGGGCAGAGATCC 60
QY 86 GlnSerProGlnGlyGlnGluGlnArgGlnHisLeuHisLeuMetValGlnLeuLeu 105
   |||||
Db 61 CAGAGTCCCCAGAAAGAGAGAGAGCAGAGCAGACCTGCACCTCATGGTACAGTGCCTG 120
QY 106 ArgProGlnAspAspIleArgLeuAlaGlnLeuGluAlaProArgProProArgLeu 125
   |||||
Db 121 AAGCCGAGAGATGACATCCGCTGGGAGCCAGACTGAGGACACCCGCGCTCCCGGCTC 180
QY 126 ArgTyrLeuLeuValValSerThrArgGlnGlyGlnGlyLeuSerGlnAspGlnThrVal 145
   |||||
Db 181 CCCTACTCTCTGTAATTTCTACACAGAGAGAGAGAGAGTCTGAGCCAGATGAGAGCGT 240
QY 146 LeuLeuGlnValAlaAspPheProAspSerSerSerProSerCysThrLeuGlyLeuValLeu 165
   |||||
Db 241 CTCCTGGGGGTGGATTCCCTGACACAGCAGCTCCCGCAGCTGCACCTGGGGCTGCTTG 300
QY 166 ProLeuTyrSerAspThrGlnValTyrLeuAspGlyAspGlyGlyPheSerValThrSer 185
   |||||
Db 301 CCCCTCTGAGTGTACACCCAGAGGTACTTAGATGAGAGAGGGGGCTTCAAGCTGAGCTGT 360
QY 186 GlyGlyGlnSerArgIlePheLeuProIleSerIleGlnThrMetTyrPalaThrLeuGln 205
   |||||
Db 361 GGTGGGCAAAAGCGGATCTTCAAGCCCATCTCCATCCAGCATGTGGGCAACATCCAG 420
QY 206 ValLeuHisGlnAlaGlyGlnAlaAlaLeuGlySerGlyLeuValProGlyGlySerAla 225
   |||||
Db 421 GTATTGTACCAAGCATGTAGGAGCTTAGGAGCGGCTTGTACCGGGGTGGCAGTGGCC 480
QY 226 LeuThrTyrPalaSerHisTyrGlnGluArgLeuAsnSerGlnGlnSerCysLeuAsnGlu 245
   |||||
Db 481 CTCACCTGGGCCAGCCACATACAGAGAGAGACTGACTCCGAAACAGAGCTGCTCAATGAG 540
QY 246 ThrPThrAlaMetAlaAspLeuGlnSerLeuArgProProSerAlaGlnProGlyGlySer 265
   |||||
Db 541 TCGAGCGGTATGCGCCAGCTGAGTCTGCGTGTCCAGCGCCGAGCGCTGGCGGTGCC 600
QY 266 SerGlnGlnGlnMetGlnAlaIleArgAlaGlnLeuTyrPalaThrLeuAspVal 285
   |||||
Db 601 TCAGAAACAGAGACAGATGAGAGAGGAGCTCGTGTGAGCTGTGGAAGTGTGCGATGTC 660
QY 286 SerAspLeuGlnSerValThrSerTyrGlnLeuArgGlnAlaLeuGlnLeuArgLeuGln 305
   |||||
Db 661 AGTGAAGCTGAGAGAGTGTACTTCCAAAGAGATCCGCCAGGCTCTGAGCTGGCGCTGGG 720
QY 305 yLeuProLeuGlnGlnTyrArgAspPheIleAspAsn--GlnMetLeuLeuValAla 324
   |||||
Db 721 GCTCCCCCTTCCAGCAGTACCGGACTTCATGCACAAACAGATTCCTGCTGCTGCTGGT 780
QY 324 IAGlnArgAspArgAlaSerArgIlePheProHisLeuTyrLeuGlySerGlnTyrPala 344
   |||||
Db 781 CACAGCGGGAGCGAAGCTCCGATC-TTCCCCCAACTCTA-GTTGGTCAGAGATGGAACG 838
QY 344 IAlaAsnLeuGlnGlnLeuGlnArgAsnArg 354
   |||||
Db 839 CAGA-AACTTGAGAGAGCTCCGAGAGACAGG 869

```

```

REFERENCE 1 (bases 1 to 1024)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LCM1905 row: b column: 01
High quality sequence stop: 655.
Location/Qualifiers
1..1024
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5431968"
/clone_lib="NIH_MGC_98"
/tissue_type="astrocytoma grade IV, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: brain; Vector: pOTB7; site_1: XhoI; site_2:
EcoRI; CDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-CDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
BASE COUNT 221 a 305 c 328 g 170 t
ORIGIN
Alignment Scores:
Pred. No.: 2.4e-127 Length: 1024
Score: 1318.00 Matches: 283
Percent Similarity: 76.74% Conservatave: 4
Best Local Similarity: 75.67% Mismatches: 16
Query Match: 54.11% Indels: 71
DB: Gaps: 4
US-09-761-640-4 (1-471) x BM560192 (1-1024)
QY 25 AlaValGlnArgArgSerArgLeuGlnArgArgGlnSerPheAlaValLeuArgGlyAla 44
   |||||
Db 3 GCGGTCCAGCGAAGAGTGTGACTCCAGCGAAGCAGACGTTTGGGTGCTCCGTGGGCT 62
QY 45 ValLeuGlyLeuGlnAspGlyGlyAspAsnAspAspAlaAlaGlnAlaSerSerGluPro 64
   |||||
Db 63 GTCTGTGGACTCGAGATGAGAGGGAGCANTGATGATGACAGAGAGGCCAGTTCTGAGCCA 122
QY 65 ThrGlnAlaProSerGluGluLeuHisGlyAspGlnThrAspPheGlyGlnGly 84
   |||||
Db 123 ACAGAGAAAGCCCCAGTAGAGAGAGCTCCAGGGGAGCAACAGACTTCGGGGAAGA 182
QY 85 SerGlnSerProGlnGlyGlnGluGlnArgGlnHisLeuHisLeuMetValGlnLeu 104
   |||||
Db 183 TCCAGAGTCCCCAGAGAGAGAGAGAGCAGAGCAGACCTGCACCTCATGTACAGCTG 242
QY 105 LeuArgProGlnAspAspIleArgLeuAlaGlnLeuGluAlaProArgProProArg 124
   |||||
Db 243 CTGAGGCCCGCAGAGATGACATCCGCTGGGAGCCAGAGCTGAGAGCAGCCCGGCTCCCGG 302
QY 125 LeuArgTyrLeuLeuValValSerThrArgGlnGlyGlnGlyLeuSerGlnAspGlnThr 144
   |||||
Db 303 CTCCTGACTCTGTAATTTCTACACAGAGAGAGAGAGAGTCTGAGCCAGATGAGAGG 362
QY 145 ValLeuLeuGlyValAlaAspPheProAspSerSerSerProSerCysThrLeuGlyLeuVal 164
   |||||
Db 363 GTCTCTCTGGGGGTGGATTTCCTGTACAGCAGCTCCCGCAGCTGACACCTGGGCTGGT 422
QY 165 LeuProLeuTyrSerAspThrGlnValTyrLeuAspGlyAspGlyGlyPheSerValThr 184

```

```

|||||
Db 423 TTGCCCCCTGTGAGTACAGCCAGGTACTGATGAGAGAGGGGCTTCAGCTGACG 482
QY 185 SerGIUGInSerArgIIlePheIysProIleSerIleGlnThrMetTrpAlaThrIleu 204
Db 483 TCTGGTGGGCAAGGCGGATGTTCTTAAGCCCATCTCCATCCAGACCATGTGG----- 533
QY 205 GlnValIleuHisGlnAlaCysGlnAlaAlaIleuGlySerGlyLeuValProGlyGlySer 224
Db 533 ----- 533
QY 225 AlaLeuThrTrpAlaSerHisIleTyrGlnGlnIleuAsnSerGlyIleuSerCysLeuAsn 244
Db 533 ----- 533
QY 245 GluTrpThrAlaMetAlaAspLeuGlnSerIleuArgProProSerAlaGluProGlyGly 264
Db 533 ----- 533
QY 265 SerSerGIUGInGlnGlnMetGlnGlnAlaIleArgAlaGlnIleuTrpIleValLeuAsp 284
Db 534 TCCCTCAGAACAGAGACATGAGAGCGCATCCGCTGTGAGCTGTGAAAGTGTGGAT 593
QY 285 ValSerAspIleuGlnSerValThrSerIleuGlnIleArgGlnAlaIleuGlnIleuArgLeu 304
Db 594 GTCACTGACCTGAGAGAGTCACTTCCCAAGAGATCCGCCAGGCTGTGAGCTGGCGCTG 653
QY 305 GlnIleuProIleuGlnGlnIleuArgAspPheIleAspAsnGlnMetLeuIleuValAla 324
Db 654 GGGTCTCCCTCCAGAGTACCGTACTTCTGACAAACCGAGTCTGCTCTGTGGTGA 713
QY 325 GlnArgAspArgAlaSerArgIIlePheProHisIleuTrpIleuGlySerGlyTrpAsnAla 344
Db 714 CAGGGGAGCGAGCGCTCCCGCATCTTCCCAACTCTTACCTGGGCTCAGATGGAACCA 773
QY 345 Ala-AsnIleuGln-GluIleuGlnIleArgAsn-ArgValThrHisIleuAsnMetAla--- 362
Db 774 GCCAAACCTGGAGGAGCTGCCAAGAAAGAGGTCCACCCATCTTGAAATGGGCGCC 833
QY 363 ArgGluIleAspAspPheTrpProGlnIleArgPheThrIleu-----HisAsnValIleuArgLeu 380
Db 834 GGGAGATTGGCAACTTCTTAACCTGAAACGCGCTTCAACCTTACCACCAATGTGGCGCTC 893
QY 381 TrpAspGlnIleuSerAlaGlnIleuLeuProHisTrp 392
Db 894 CTGGGGAATGAGGA-----ATCGGCCCCAGCTGG 923

RESULT 5
Bg750408 812 bp mRNA linear EST 15-MAY-2001
LOCUS 602709243F1 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:4845707 5',
ACCESSION Bg750408
VERSION Bg750408.1 GI:14061061
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 812)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L1CM1685 Row: f Column: 12

```

```

FEATURES
    source
        1..812
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:4845707"
            /clone_lib="NIH_MGC_43"
            /tissue_type="normal pigmented retinal epithelium"
            /lab_host="DH10B (phage-resistant)"
            /note="Organ: eye; Vector: pOTB7; Site.1: XhoI; Site.2:
            EcoRI; cDNA made by oligo-dT priming. Directionally
            cloned into EcoRI/XhoI sites using the following 5'
            adaptor: GGCACGAG(G). Library constructed by Ling Hong
            in the laboratory of Gerald M. Rubin (University of
            California, Berkeley) using ZAP-cDNA synthesis kit
            (Stratagene) and Superscript II RT (Life Technologies).
            Note: this is a NIH_MGC library. |"

BASE COUNT 167 a 250 c 254 g 141 t
ORIGIN

Alignment Scores:
Pred. NO.: 8.76e-126 Length: 812
Score: 1301.50 Matches: 264
Percent Similarity: 97.06% Conservative: 0
Best Local Similarity: 97.06% Mismatches: 7
Query Match: 53.43% Indels: 3
DB: 12 Gaps: 1

US-09-761-640-4 (1-471) x Bg750408 (1-812)
QY 66 GlnIleuAlaProSerGlnGlnGlnIleuHisGlyAspGlnThrAspPheGlyGlnGlySer 85
Db 1 CGAAGGCGCCCGAGTACAGAGAGAGCTCCAGGGGACCGACAGACTTGGGCAAGATCC 60
QY 86 GlnSerProGlnIleuGlnGlnIleuArgGlnHisIleuHisIleuMetValGlnIleu 105
Db 61 CAGAGTCCCGCAAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
QY 106 ArgProGlnAspAspIleArgLeuAlaAlaGlnIleuGlnAlaProArgProProArgLeu 125
Db 121 AGCGCGCAGATGACATCGCTGGCAGCCAGCGAGGAGACCCCGCGCTCCCGCGCTC 180
QY 126 ArgTrpIleuValValSerThrArgGlnGlnGlyIleuSerGlnAspGlnThrVal 145
Db 181 CGTATCTCGCTGAGTATTCTTACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
QY 146 LeuIleuGlnValAspPheProAspSerSerSerProSerCysThrLeuGlyLeuValLeu 165
Db 241 CTCCGCGCGCTGGATTTCCCTGACAGCAGCTCCCGCACCTGCACCCCTGGCCTGTTG 300
QY 166 ProIleuTrpSerAspPheGlnValIleuAspGlyAspGlyIleuPheSerValThrSer 185
Db 301 CCCCTCTGAGTACACCCAGGTGACTTACATGAGAGAGGGGGCTTACGCTGACGCTC 360
QY 186 GlnGlyGlnSerArgIIlePheIysProIleSerIleGlnThrMetTrpAlaThrIleuGln 205
Db 361 GGTGGCAACACCGGATCTTCAAGCCCATCTCCATCCAGACCATGTGGCCACACTCCAG 420
QY 206 ValIleuHisGlnAlaCysGlnAlaAlaIleuGlySerGlyLeuValProGlyIleuVal 225
Db 421 GTATTGGCACCAAGCATGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
QY 226 LeuThrTrpAlaSerHisIleTyrGlnGlnIleuAsnSerGlyIleuSerCysLeuAsnGln 245
Db 481 CTCACCTGGGCGAGCCACTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
QY 246 TrpThrAlaMetAlaAspLeuGlnSerIleuArgProProSerAlaGluProGlyGlySer 265
Db 541 TGGACGGGTATAGCGGAGCTGTGGGCTCCACAGCCCGGAGCTGGCGGGTCC 600
QY 266 SerGIUGInGlnGlnMetGlnGlnAlaIleArgAlaGlnIleuTrpIleValLeuAspVal 285
Db 601 TCAGAACAGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660

```

Oy 286 SerAspIeuGlSerValThrSerIysGluIleArgGlnAlaIeuGluIeuArgIeuGly 305
|||||
Db 661 AGTGAACCTGAGAGTGTACTCTCCAAGAGATCCGACGCTGTGGAGCTTGCGCTGGG 720
Oy 306 LeuProIeuGlnGlnTyrArgAspPheIleAspAsnGlnMetIeuIeuValAlaGln 325
|||||
Db 721 TGTCCCTCCACAGTACGCTGACTTCATCGACACAG-ATGCTGCTGCTGGTGGCAG 779
Oy 326 ArgAspArgAlaSerArgIlePheProHisLeuTyr 337
|||||
Db 780 CGGAA-CGAGCTCCCGATTTCGCCA---CTCTAC 811
RESULT 6 790 bp mRNA linear EST 27-FEB-2001
LOCUS BG326709
DEFINITION 602425578F1 NIH_MGC_14 Homo sapiens cDNA clone IMAGE:4563595 5',
mRNA sequence.
ACCESSION BG326709
VERSION BG326709.1 GI:13133146
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 790)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: DCTD/DRP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L10C1277 row: K column: 20
High quality sequence stop: 737.
Location/Qualifiers
1..790
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4563595"
/clone_lib="NIH_MGC_14"
/tissue_type="renal cell adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: Kidney; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using Zap-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 162 a 237 c 268 g 123 t
ORIGIN
Alignment Scores:
Pred. No.: 2.27e-115 Length: 790
Score: 1202.00 Matches: 236
Percent Similarity: 96.75% Conservative: 2
Best Local Similarity: 95.93% Mismatches: 7
Query Match: 49.34% Indels: 1
DB: 12 Gaps: 0
US-09-761-640-4 (1-471) x BG326709 (1-790)
Oy 1 MetalaleuValThrValSerArgSerProGlySerGlyAlaSerThrProValGly 20
|||||
Db 50 ATGGCCCTGTGACAGTAGCCGTTGCCCCCGGGCAGCGGCGCTCCACGCCCGTGCGG 109

Oy 21 ProTPAspGlnAlaValGlnArgSerArgIeuGlnArgGlnSerPheAlaVal 40
|||||
Db 110 CCTGGACACAGCGGCTCCAGGAAGAGTGCATCTCCAGGAAGGACAGCTTTGGGGT 169
Oy 41 LeuArgGlyAlaValIeuGlnIleuGlnAspGlyGlyAspAsnAspAlaIaGlnAla 60
|||||
Db 170 CTCCTGGGGGCTCTCTGGAGTGCAGATGAGGGGACAAAGATGATCCACAGAGGCC 229
Oy 61 SerSerGluProThrGlnIysAlaProSerGluGluGlnIleHisGlyAspGlnThrAsp 80
|||||
Db 230 AGTTCTGACCCACAGAGAGAGGCCCGGAGTGAAGAGAGCTCCAGGGACACAGACAGC 289
Oy 81 PheGlyGlnGlySerGlnSerProGlnIysGlnIleuGlnArgGlnHisIleuHisIeu 100
|||||
Db 290 TTCGGGCAAGGATCCAGAGTCCCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 349
Oy 101 MetValGlnIleuIeuArgProGlnAspAspIleArgIeuAlaIaGlnIleuAlaPro 120
|||||
Db 350 ATGGTACACACTCTGAGGCGCAGAGTGAATCCGCTGGCAGCCAGCTGAGAGGACCC 409
Oy 121 ArgProProArgIeuArgTyrIleuIeuValIaSerThrArgGluGlyGluIeuSer 140
|||||
Db 410 CGGCTCCCGGCTCCGCTACCTGCTGATTTCTACACGAGAGAGAGAGAGTCTGAGC 469
Oy 141 GlnAspGluThrValIleuIeuGlyValAlaAspPheProAspSerSerProSerCysThr 160
|||||
Db 470 CAGGATGAGACAGGTCTCTGGGCGGATTTCCCTGACAGAGAGCTCCCGCAGCTGACCC 529
Oy 161 LeuGlyIleuValIeuProIeuTrpSerAspThrGlnValTyrIleuAspGlyAspGly 180
|||||
Db 530 CTGGGCTGGTCTGCTCCCTGAGGTGACACCGGTGACTTGAATGAGACAGGGGGC 589
Oy 181 PheSerValThrSerGlyIleuSerArgIlePheIleProIleSerIleGlnThrMet 200
|||||
Db 590 TTCACGCTACGCTCTGCTGGGCAAAACCGGATCTTCAACCCATCTCCATCCAGACAG 649
Oy 201 TrpAlaThrIleuGlnValIleuHisGln-AlaCysGlnAlaIaIeuGlySerGlyIeuVa 220
|||||
Db 650 TGGGCGACACTCCAGGTATTCACCAAGCATGTAGGAGGCTTAGGACAGAGCTGGT 709
Oy 220 ProGlyIleuSerAlaIeuThrTrpAlaSerHisIleuGlnIleuArgIleuSerGlu 240
|||||
Db 710 ACCGGGTGGCACTGCTTCACTGGGCAAGCATACAGAGAGAGACTGCACTCCGAACA 769
Oy 240 nserCysIeuAsnGlu 245
|||||
Db 770 GAACCTGCTCATATGAG 785
RESULT 7 846 bp mRNA linear EST 15-AUG-2000
LOCUS BE563259
DEFINITION 601335565F1 NIH_MGC_39 Homo sapiens cDNA clone IMAGE:3689369 5',
mRNA sequence.
ACCESSION BE563259
VERSION BE563259.1 GI:9807071
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 846)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

Plate: L1CM384 row: a column: 18
High quality sequence stop: 709.

FEATURES

source

Location/Qualifiers
1..846

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3689369"
/clone_1lb="NIH_MGC_39"
/tissue_type="adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA Synthesis Kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 172 a 252 c 266 g 156 t
ORIGIN

Alignment Scores:

Pred. No.: 1.32e-109 Length: 846
Score: 1147.50 Matches: 251
Percent Similarity: 92.96% Conservative: 0
Best Local Similarity: 92.96% Mismatches: 17
Query Match: 47.11% Indels: 8
DB: 10 Gaps: 1

US-09-761-640-4 (1-471) x BE563259 (1-846)

Oy 128 LeuLeuValValSerThrArgGluGlyGluSerGlnAspGluThrValLeuLeu 147
|||||
Db 2 CTGCTGGTAGTTCTACACGACAGAGAGGTCTGAGCCAGATGAGCGTCTCTCG 61
148 G1yValAspPheProAspSerSerSerProSerCysThrLeuGlyLeuValLeuProLeu 167
62 GGCGGTGATTCCTCGACAGCAGCTCCCGACCTGACCTGGCGTGTGCCCCCTC 121
Oy 168 TTPSeraPThrGlnValTyrLeuAspGlyAspGlyGlyPheSerValThrSerGlyGly 187
122 TGGAGTGCACCCAGCTGACTTACAGAGAGCGGGGCTTACGGTGCAGCTGCTGG 181
Oy 188 GlnSerArgIlePheLysProIleSerIleGlnThrMetTrpAlaThrLeuGlnValLeu 207
182 CAAAGCCGATTTCAAGCCCTCTCCATCCAGACCATGTGGCCACACTCCAGGTATTG 241
Oy 208 HisGlnAlaCysGluAlaAlaLeuGlySerGlyLeuValProGlyGlySerAlaLeuThr 227
242 CACCAAGCATGTGAGAGGAGCTCTAGGACGGCGCTTGTACGGGTGGCAGTGCCCTCAC 301
Oy 228 TrpAlaSerHisTyrGlnGluArgLeuAsnSerGluGlnSerCysLeuAsnGluTrpThr 247
302 TGGGCGAGCCTTACAGAGAGAGTGAATCCGAACAGCTGCCTCAATGATGAGGAGC 361
Oy 248 AlaMetAlaAspLeuGluSerLeuArgProProSerAlaGluProGlyGlySerSerGlu 267
362 GGTATGGCCGACCTGGAGTCTCTGCGGCTCCAGCGCCGAGCTGGCGGCTCTCAGAA 421
Oy 268 GlnGlnGlnMetGluGlnAlaIleArgAlaGluLeuTrpLysValLeuAspValSerAsp 287
422 CAGGAGCAGATGAGACAGGCGATCCGTGCTGAGTGTGAAGTGTGATGTCAGTGAC 481
Oy 288 LeuGluSerValThrSerLysGluIleArgGlnAlaLeuGluLeuArgLeuGlyLeuPro 307
482 GTGGAGAGTGTACTTCCAAATAGATCCG-CAGGCTGTGGAGCTCGGCTGGGCTCCCC 540
Oy 308 LeuGlnGlnTyrArgAspPheIleAspAsnGlnMetLeuLeuValAlaGlnIleArgAsp 327
541 CTCACGACGTACCGTACTTCTACACAAAG-ATGCTGCTCTGTGTGGCACAAGCGGAGC 599
Oy 328 ArgAlaSerArgIlePheProHisLeuTyrLeuGlySerGluTrpAsnAlaIleAsnLeu 347
|||||

Db 600 CGAGCTCCCGCATCTTCCCGCACTCTACTG-CGCTCAGATGGAGCAGCAACACCTG 658
Oy 348 GlnGluLeuGlnArgAsnArgValThrHisIleLeuAsnMetAlaArgGluIleAspAsn 367
659 GAGAGCTGACAGAGACAGGAGGTACACCATTTGACATGCC-CGGAGATGTA-AAC 716
Oy 368 PheTyrProGluArgPheThrTyrHisAsnValArgLeuTrpAspGluGlnSerAlaGln 387
717 TTCTAACCTGAGCGGCTTACC-TACCACATGTGGGCTTGTGGATGAGAGATCGGGCCAC 775
Oy 388 LeuLeuProHisTrpLysGluThrHisArg 397
776 TGTGGCCCT-----GTAGGAGCAGCACCT 799

RESULT 8

LOCUS

DEFINITION

ACCESSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

Location/Qualifiers
1..910

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4424070"
/clone_1lb="NIH_MGC_89"
/tissue_type="hypernephroma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: kidney; Vector: PCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.3 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

BASE COUNT

189 a 270 c 304 g 146 t 1 others

ORIGIN

Alignment Scores:
Pred. No.: 1.49e-109 Length: 910
Score: 1147.50 Matches: 246
Percent Similarity: 92.91% Conservative: 3
Best Local Similarity: 91.79% Mismatches: 15
Query Match: 47.11% Indels: 8
DB: 12 Gaps: 2

US-09-761-640-4 (1-471) x BG169297 (1-910)

Oy 179 G1yGlyPheSerValThrSerGlyGlyGlnSerArgIlePheLysProIleSerIleGln 198
Db 2 GGGGGCTTCAAGGTGACGCTGTGTGGGCAAGCGGATTTCAAGCCCATCTCCATCCAG 61
Oy 199 ThrMetTrpAlaThrLeuGlnValLeuHisGlnAlaCysGluAlaIleLeuGlySerGly 218
|||||

Db 62 ACCATGTGGGCCACTCCAGGTATG-CACCAAGCATGTGAGGACAGCTTAGCAGCGC 120
 QY 219 LeuValProGlyGlySerAlaLeuThrTrpAlaSerHisTyrGlnGluArgLeuAsnSer 238
 Db 121 CTTGTACCGGGTGGAGCTGCTCACCCTGAGCCAGCCACTACCAAGAGACTGAACCTCC 180
 QY 239 GluGlnSerCysLeuAsnGluTrpThrAlaMetAlaAspLeuGlnSerLeuArgPro 258
 Db 181 GAACAGAGCTGCTCATATGATGAGCGGCTATGGCCGACTGAGTCTCTGGCGCTCC 240
 QY 259 SerAlaGluProGlyGlySerSerGluGlnGluMetGlnAlaIleArgAlaGlu 278
 Db 241 AGCGCCGAGCTGGCGGCTCTCAGACAGAGACAGATGAGCAGCGATCCCTGCTGAG 300
 QY 279 LeuTrpValValLeuAspValSerAspLeuGlnSerValThrSerLysGluIleArgGln 298
 Db 301 CTGTGAAAGTGTGGATGTCACTGACTGAGAGTGTCTCTCCAAAGAGATCCG-CAG 359
 QY 299 AlaLeuGluLeuArgLeuGlnLeuProLeuGlnGlnTyrArgAspPheIleAspAsnGln 318
 Db 360 GCTCTGAGCTGGCGCTGGGGGCTCCCTCCAGCAGTACCGTACCTTCATCGACCAACAG 419
 QY 319 MetLeuLeuLeuValAlaGlnArgAspArgAlaSerArgIlePheProHisLeuTyrLeu 338
 Db 420 ATGCTGCTGCTGTGGCAGACGCGGACCGAGCTCCCGCATCTTCCCAACCTCTACCTG 479
 QY 339 GlySerGluTrpAsnAlaAlaAsnLeuGlnGluLeuGlnArgAsnArgValThrHisIle 358
 Db 480 GGTCTGAGTGGAGCCAGCAACCTGGAGAGCTGCAGAGGAACAGGCTACCCACATC 539
 QY 359 LeuAsnMetAlaArgGluIleAspAsnPheTyrProGluArgPheThrTyrHisAsnVal 378
 Db 540 TTGAAATGATGCGCGGAGATNGACAACTTCACTGAGGCGCTTCACTACCAATGTG 599
 QY 379 ArgLeuTrpAspGlnLeuSerAlaGlnLeuLeuProHisTyrLysGluThrHisArgPhe 398
 Db 600 CGCTC-TGGATGTAGAGTGGCCAGCTGCTGC--GACTGGAAGAGAGACCGCGTTTC 655
 QY 399 IleGlu-AlaAlaArgAlaGlnGlnTyrHisValLeuValHisCysLysMetGlyValSe 418
 Db 656 ATGAGAGGCTGCAGAGACAGCAGGACCCAGCTGCTGCTCACTGCAAGATGGCGCTCAG 715
 QY 418 ArgSerAlaAlaThrValLeuAlaTyrAlaMetLysGln-TyrGluCysSerLeuGlnG 438
 Db 716 CCGGTACGCGGACAGAG--GCTGGTATGTCATGAAGCAAGTACGAGGCTGGAGC 773
 QY 438 ImlAlaLeuArgHisValGln 444
 Db 774 GGCT--GGGCACTGTGAG 790
 RESULT 9
 LOCUS BE869222 726 bp mRNA linear EST 20-OCT-2000
 DEFINITION 60144503orf1 NIH_MGC_65 Homo sapiens cDNA clone IMAGE:3849342 5',
 mRNA sequence.
 ACCESSION BE869222
 VERSION BE869222.1 GI:10317998
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 726)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 plate: L1AM9566 row: k column: 07
 High quality sequence stop: 607.
 Location/Qualifiers
 FEATURES
 source
 1..726
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3849342"
 /clone_lib="NIH_MGC_65"
 /tissue_type="adenocarcinoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: colon; Vector: pCMV-SPORT6; Site: 1: NotI;
 Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.8 kb. Library constructed by Life
 Technologies."

BASE COUNT 137 a 235 c 226 g 128 t
 ORIGIN

Alignment Scores:
 Pred. No.: 2,11e-109 Length: 726
 Score: 1144.50 Matches: 230
 Percent Similarity: 96.69% Conservative: 4
 Best Local Similarity: 95.04% Mismatches: 7
 Query Match: 46.98% Indels: 2
 DB: 12 Gaps: 1

US-09-761-640-4 (1-471) x BE869222 (1-726)
 QY 93 GluGlnArgGlnHisLeuHisLeuMetValGlnLeuLeuArgProGlnAspAspIleArg 112
 Db 3 GACAGAGGAGGAGCAGCTGCACCTCATGTAGTACACTGCTGAGCGCGAGATGACATCCC 62
 QY 113 LeuAlaAlaGlnLeuGlnAlaProArgProArgLeuArgTyrLeuValValSer 132
 Db 63 CTGGCAGCCAGCTGAGGAGGACCCCGGCTCCCGGCTCCGCTACTGCTGTGATTTCT 122
 QY 133 ThrArgGlnGlyGlnGlyLeuSerGlnAspGluThrValLeuLeuGlyValAspPhePro 152
 Db 123 ACAGAGAAAGAGAGAGTGTGAGCCAGAGATGAGACGAGTCTCTGGCGGTATTCCTCC 182
 QY 153 AspSerSerSerProSerCysThrLeuGlyLeuValLeuProLeuThrPheSerPheGln 172
 Db 183 GACAGAGCTGCCCCAGCTGCACCTGGCGGTGTCTTGCCTCTGGAGTGCACACCCAG 242
 QY 173 ValTyrLeuAspGlyAspGlyGlyPheSerValThrSerGlyValGlnSerArgIlePhe 192
 Db 243 GTTACTTATGATGAGACAGGAGGCTTACGCGTACGCTGTGTTGGCAAGCCGATCTTC 302
 QY 193 LysProIleSerIleGlnThrMetTrpAlaThrLeuGlnValLeuHisGlnAlaCysGln 212
 Db 303 AAGCCATCTCCATCCAGACCATGTGGCCACACCTCCAGTATTCACCAAGCATGTGAG 362
 QY 213 AlaAlaLeuGlySerGlyLeuValProGlyGlySerAlaLeuThrTrpAlaSerHisTyr 232
 Db 363 GCAGCTCTAGGACGCGCTTGTACCGGGGTGCGAGTCCCTCACTGGGCGACCCACTAC 422
 QY 233 GlnGluArgLeuAsnSerGlnGlnSerCysLeuAsnGluTrpThrAlaMetAlaAspLeu 252
 Db 423 CAGGAGAGACTGAACTCCGACAGAGACTGCTCAATGATGGAGCGCTATGGCCGACTG 482
 QY 483 GACTCTCTGGGCTCCACAGCGCGGAGCTGGGGGTCTCCAGAACAGAGCAGATGAGAG 542
 Db 273 GlnAlaIleArgAlaGluLeuTrpLysValLeuAspValSerAspLeuGlnSerValThr 292
 QY 543 CAGCGCATCCGTGCTGAGCTGTGCAAGTGTGCGATGTAGTACCTGGAGATGTCTACT 602
 Db 293 SerLysGlnIleArgGlnAlaLeuGlnGluLeuArgLeuGlyLeuProLeuGlnGlnTyrArg 312
 Db 603 TCCAAAGAGATCCGCCAGGCTCTGAGACTGCGCTGGGGCTCCCTCCCGCCAGTACCGC 662

QY 313 AspHeileaspasngImetleuleuleuValAlaInArgAspArgAlaSerArgIle 332
 Db 663 GACTTCAATCGACACAG-ATGCTGCTGTGTGTGGCCAGC---GGGCCGAGCTCGGCTC 718
 QY 333 PhePro 334
 Db 719 TTCCCC 724

RESULT 10
 BE737087 919 bp mRNA linear EST 15-SEP-2000
 LOCUS 60130485F1 NIH_MGC_39 Homo sapiens cDNA clone IMAGE:3639229 5',
 DEFINITION mRNA sequence.
 ACCESSION BE737087
 VERSION BE737087.1 GI:10151079
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE 1 (bases 1 to 919)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: ATCC

CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LHC341 row: h column: 14
 High quality sequence start: 62
 High quality sequence stop: 769.
 Location/Qualifiers

FEATURES

source

1. 919
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3639229"
 /clone_id="NIH_MGC_39"
 /tissue_type="adenocarcinoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: pancreas; Vector: pOTB; Site: 1: XhoI;
 Site: 2: EcoRI; CDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACGAG(G). Library constructed
 by Ling Hong in the laboratory of Gerald M. Rubin
 (University of California, Berkeley) using ZAP-CDNA
 synthesis kit (Stratagene) and Superscript II RT (Life
 Technologies)."
 BASE COUNT 170 a 269 c 316 g 164 t
 ORIGIN

Alignment Scores:

Pred. No.: 1.13e-103 Length: 919
 Score: 1091.50 Matches: 236
 Percent Similarity: 91.29% Conservative: 5
 Best Local Similarity: 89.39% Mismatches: 18
 Query Match: 44.81% Indels: 8
 Gaps: 12

US-09-761-640-4 (1-471) x BE737087 (1-919)

QY 175 leuaspGlyaspGlyGlyPheSer-ValThrSerGlyGlyInserArgIlePheLeu 194
 Db 42 TTATATGAGACGGGGT---TCAGGTGACGCTGTGGCAACCGGATCTTCAAGC 98
 QY 194 GATCCTCATCGACACCATGTCGACACTCCAGTATG-CACCAACGATGTGAGCAGC 157
 Db 99 CATCTCATCGACACCATGTCGACACTCCAGTATG-CACCAACGATGTGAGCAGC 157

QY 214 aleuGlySerGlyLeuValProGlyGlySerAlaLeuThrTrpAlaSerHisTyrgInG 234
 Db 158 TCTAGGACAGCGGCTTGTATACCGGGTGGAGTCCCTCATCGTGGGCGACCACTACACAGA 217
 QY 224 uArgLeuAsnSerGlyInserCysLeuAsnGluTrpPhrAlaMetAlaAspLeuG 254
 Db 218 GAGACTGAACCTCCGAAACAGACAGCTGCTCAATGAGGAGGAGGCTATGGCGGACCTGAGT 277
 QY 254 rLeuArgProProSerAlaGluProGlyGlySerSerGlyInGInGInGInGInG 274
 Db 278 TCTGGGCTCTCCAGCGCGGAGCTGGCGGCTCTCAAGAACAGGAGCAGATGAGCGGGC 337
 QY 274 aileArgAlaGluLeuTrpValLeuAspValSerAspLeuGInGInGInGInG 294
 Db 338 GATCCGCTCTGAGCTGTGGAAGTGTGGATGTCGTGACCTGGAGATGTGCTTCAG 397
 QY 294 sGluIleArgGlnAlaLeuGluLeuArgLeuGlyLeuProLeuGInGInGInG 314
 Db 398 AGAGATCCGCGAGGCTCTGGAGCTGCGCTGGGCTCCCTCCAGCAGTACCGTACTT 457
 QY 314 eileaspasngImetleuleuleuValAlaInArgAspArgAlaSerArgIlePhe 334
 Db 458 CATCGACACAGATGCTGTGTGTGTGGCACAGCGGAGCGAGCTCCGATCTTCC 517
 QY 334 oHisLeuTrpLeuGlySerGluTrpAsnAlaAlaAsnLeuGluGluLeuGlnArgAsn 354
 Db 518 CCACCTTACCTGGGCTGAGAGTGAAGCGCAACCTGGAGAGCTCCGAGAGGACAG 577
 QY 354 gValThrHisIleLeu-AsnMetAlaArgGluIleAspAsnPheTyrrProGluArgPhe 374
 Db 578 GGTCAACCACTCTGTAACTAGGCGCGGAGATGACAA-TTCAACCTGAGCCCTTCA 636
 QY 374 hrTyrrHis-AsnValArgLeuTrpAspGluGInGInGInGInGInGInG 393
 Db 637 CTTACACAGATGCTGCTGTGGATGAGAGTGGC-CAGTCTCTCCGACCTGAGG 695
 QY 394 GluThrHisArgPheIleGluAlaAlaArgAlaGlnGlyThrHisValLeuValHisCys 413
 Db 696 GAGACGACACCGCTTATGAGGCTGCGAGAGACAGCGGACCCACGCTGCTGCCCTGG 755
 QY 414 -LysMetGlyValSerArgSerAlaAlaThrValLeuAlaTrpAlaMetLysGlnTyrg 433
 Db 756 CAAGATGGGGGCTGACGCGTCAAGGGGCGCGGCTGTATGCGGTGAAGACGTACGA 815
 QY 433 uCys 434
 Db 816 ATGC 819

RESULT 11
 LOCUS BM563401 943 bp mRNA linear EST 20-FEB-2002
 DEFINITION AGNCOCORT_6589387 NIH_MGC_98 Homo sapiens cDNA clone IMAGE:5441645
 5', mRNA sequence.
 ACCESSION BM563401
 VERSION BM563401.1 GI:18810292
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE 1 (bases 1 to 943)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov

Plate: L1CM1916 row: e column: 06
High quality sequence stop: 618.
Location/Qualifiers
1. 943

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5441645"
/clone_lib="NIH_MGC_98"
/tissue_type="astrocytoma grade IV, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: brain; Vector: pOTB7; Site:1: XhoI; Site:2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

BASE COUNT 201 a 284 c 305 g 153 t
ORIGIN

Alignment Scores:

Pred. No.: 1.91e-103 Length: 943
Score: 1089.50 Matches: 236
Percent Similarity: 75.24% Conservatave: 4
Best Local Similarity: 73.98% Mismatches: 9
Query Match: 44.72% Indels: 70
DB: 13 Gaps: 3

US-09-761-640-4 (1-471) x BM563401 (1-943)

QY 25 AAlaValGlnArgSerArgLeuGlnArgGlnSerPheAlaValLeuArgGlyAla 44
DB 27 GGGGTCACAGCGAAGAGTGATGACTCCAGCGAAGGAGAGCTTGCGGTGCTCCGTGGGGT 86
QY 45 ValLeuGlyLeuGlnAspGlyGlyAspAsnAspAspAlaAlaGlnAlaSerSerGluPro 64
DB 87 GTCCCTGGGACCTGACAGATGAGGAGGACACATGATGACGACGAGGCCAGTTCGAGCCA 146
QY 65 ThrGluValAspSerGlnGlnGlnLeuHisGlyAspGlnThrAspPheGlyGlnGly 84
DB 147 ACAGAGAAAGGCCCGACAGTACAGAGAGCTCCAGGGGACCAACAGACTTCGGGCAAGGA 206
QY 85 SerGlnSerProGlnLysGlnGlnGlnArgGlnHisLeuHisLeuMetValGlnLeu 104
DB 207 TCCGACAGATCCCAAGAGAGAGAGAGAGGAGGACGACCTGACCTCATGTGACAGCTG 266
QY 105 LeuArgProGlnAspAspLeuArgLeuAlaAlaGlnLeuAlaProArgProArg 124
DB 267 CTGAGGCGGACAGATGACATCCGCTGCGACGCCAGCTGGAGGACCCCGGCTCCCGCG 326
QY 125 LeuArgTyrLeuLeuValValSerThrArgGlnGlnGlnGlnLeuSerGlnAspGluThr 144
DB 327 CTCGGCTACCTGCTGTTGTTCTACACAGAGAGAGAGAGTCTGAGCCAGATGAGACG 386
QY 145 ValLeuLeuGlyValAspPheProAspSerSerProSerCysThrLeuGlyLeuVal 164
DB 387 GTCCCTCTGGGGGTGATTTCCCTGACAGACGCTCCCGACGTGACCTGGGCGCTGCTC 446
QY 165 LeuProLeuTrpSerAspThrGlnValTyrLeuAspGlyAspGlyGlySerValThr 184
DB 447 TTGCCCCCTGGAGTACACCCAGGTTGACTTGAATGAGAGGGGGGCTTCAGCGTGGAG 506
QY 185 SerGlyGlyGlnSerArgIlePheLysProLieserIleGlnThrMetTrpAlaThrLeu 204
DB 507 TCTGGTGGCAACCCGCGATCTTCAAGCCCATCTCATCCAGACCATGTGG----- 557
QY 205 GlnValLeuHisGlnAlaCysGlnAlaAlaLeuGlySerGlyLeuValProGlyGlySer 224
DB 557 ----- 557
QY 225 AlaLeuThrTrpAlaSerHisTyrGlnGlnGlnArgLeuAsnSerGlnGlnSerCysLeuAsn 244

DB 557 ----- 557
QY 245 GluTrpThrAlaMetAlaAspLeuGlnSerLeuArgProProSerAlaGluProGlyGly 264
DB 557 ----- 557
QY 265 SerSerGlnGlnGlnGlnMetGlnGlnAlaIleArgAlaGluLeuTrpLysValLeuAsp 284
DB 558 TCTCTGCAACACAGACAGTACAGTACAGAGGAGGATCCGCTGCTGAGACTGTGGAAT 617
QY 285 ValSerAspLeuGlnSerValThrSerLysGluIleArgGlnAlaLeuGlnLeuArgLeu 304
DB 618 GTGAGTACCTGAGAGTGTCACTTCCCAAGAGATCCCGACGCTCTGGAGCTGGCGCTG 677
QY 305 GlyLeu-ProLeuGlnGlnTyrArg-AspPheIleAsp-AsnGlnMetLeuLeuLeu--- 322
DB 678 GGGCTCCCTCCACAGATACCGGAGTATGAGAACACCAATGGCTGCTGCTGG 737
QY 323 ValAlaGln-----ArgAspArgAlaSerArgIlePheProHisLeu 336
DB 738 GTGGCACAGCCGGGAGACCCGACCTCCCGGCAATCCTTCCCGCCACCTT 788

RESULT 12
BG340220 932 bp mRNA linear EST 27-FEB-2001
LOCUS BG340220
DEFINITION 602438408F1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4556401 5', mRNA sequence.
ACCESSION BG340220
VERSION BG340220.1 GI:13146647
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 932)
NIH-MGC <http://mgc.nci.nih.gov/>.
TITILE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
CONTACT: Robert Strausberg, Ph.D.
COMMENT Email: cgap@b-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

Plate: L1CM1258 row: p column: 02
High quality sequence stop: 630.
Location/Qualifiers

FEATURES
source

1. 932
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4556401"
/clone_lib="NIH_MGC_46"
/tissue_type="leiomyosarcoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: uterus; Vector: pOTB7; Site:1: XhoI; Site:2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

BASE COUNT 160 a 284 c 326 g 162 t
ORIGIN

Alignment Scores:
Pred. No.: 5.56e-103 Length: 932
Score: 1085.00 Matches: 236
Percent Similarity: 90.98% Conservatave: 6

Best Local Similarity: 88.72% Mismatches: 21
 Query Match: 44.54% Indels: 7
 DB: 12 Gaps: 1

US-09-761-640-4 (1-471) x BG340220 (1-932)

QY 1 MetalaleuValThrValSerArgSerProProGlySerGlyAlaSerThrProValGly 20
 DB 43 ATGGCCCTGGTACAGTAGCGGTGCGCCCGGCGGCGGCGGCTCCACGCGCGTGGGG 102
 QY 21 ProTrpAspGlnAlaValGlnArgSerArgLeuGlnAlaArgGlnSerPheAlaVal 40
 DB 103 CCTGGGACCAAGCGGCTCCAGGAGGAGTGCATCCAGCAGGACAGAGCTTTCGGTG 162
 QY 41 LeuArgGlyAlaValLeuGlyLeuGlnAspGlyValAspAsnAspAlaAlaGlnAla 60
 DB 163 CTCCGCGGGCTGCTCTGGACTGAGAGTGGAGGAGCAATGATGATGACACAGAGCC 222
 QY 61 SerSerGluProThrGluLysAlaProSerGluGluLeuHisGlyAspGlnThrAsp 80
 DB 223 AGTTCTGAGCCACAGAGAGAGGCCCGGAGTGAAGAGAGCTCCAGGGGACAGACAGC 282
 QY 81 PheGlyGlnGlySerGlnSerProGlnLysGlnGlnGlnAlaArgGlnHisLeuHisLeu 100
 DB 283 TTCGGGCAAGGATCCCAAGATCCCAAGAGCAGAGGAGGAGGAGCAGCAGCTCCAGCTC 342
 QY 101 MetValGlnLeuLeuArgProGlnAspAlaArgLeuAlaAlaGlnLeuGlnAlaPro 120
 DB 343 ATGGTACAGCTGCTGAGCGCGAGGATGACATCCGCTGGCAGCCAGCTGGAGGACACC 402
 QY 121 ArgProArgLeuArgTyrLeuLeuValValSerThrArgGlnGlyLeuLeuSer 140
 DB 403 CGGCT-CCCGCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 461
 QY 141 GlnAspGluThrValLeuLeuGlyValAlaSpheProAspSerSerSerProSerCysThr 160
 DB 462 CAGGTGAGAGCGTCTCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 521
 QY 161 LeuGlyLeuValLeuProLeuThrSerAspThrGlnValTyrLeuAspGlyValSpGly 180
 DB 522 CTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 581
 QY 181 PheSerValThrSerGlyGlyGlnSerArgLeuPheLysProLysSerThrLeuThrMet 200
 DB 582 TTCACCGTAGAGTGTGGTGGGCAAGCCGAGTCTCAAGCCATC-TCCATCAAGACATG 640
 QY 201 TrpAlaThrLeuGln-ValLeuHisGlnAlaCysGlnAlaAlaLeuGlySerGlyLeu-V 220
 DB 641 TGGGCAACACTCCAGGATGATGACCCAGCATGTGAGGACAGCTTAAGCGCGGCTTTG 700
 QY 220 aLProGlyLysSerAlaLeuThrTrpAlaSerHisTyrGlnGluArgLeuAsnSerGluG 240
 DB 701 TTCCCTGGGGAGTCCCTCAGCTGGGAGCACT-TCCCGGGGGAGGCTGAGCTCCGA-C 758
 QY 240 InSerCysLeuAsnGluTrpThrAlaMetAlaAspLeuGlnSerLeuArgProProSerA 260
 DB 759 AGAGTGCTTCATGAGAGTGGCGGATATCCGAGCTGAGAGCTTGTGGG---CTCCACG 815
 QY 260 IaGluProGlyGly 264
 DB 816 CCGACGCTGGGGC 829

RESULT 13
 BE873337 897 bp mRNA linear EST 20-OCT-2000
 LOCUS BE873337
 DEFINITION 601450290F1 NIH_MGC_65 Homo sapiens cDNA clone IMAGE:3854158 5',
 mRNA sequence.
 ACCESSION BE873337
 VERSION BE873337.1 GI:10322113
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 897)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: L1AM9579 row: C column: 23
 High quality sequence stop: 632.
 Location/Qualifiers

FEATURES
 source
 1..897
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3854158"
 /clone_lib="NIH_MGC_65"
 /tissue_type="adenoecarcinoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: colon; Vector: pCMV-SPORT6; Site: 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.8 kb. Library constructed by Life
 Technologies."

BASE COUNT 168 a 254 c 322 g 152 t 1 others
 ORIGIN

Alignment Scores:
 Pred. No.: 1,49e-98 Length: 897
 Score: 1042.50 Matches: 222
 Percent Similarity: 95.28% Conservative: 0
 Best Local Similarity: 95.28% Mismatches: 10
 Query Match: 42.80% Indels: 5
 DB: 12 Gaps: 1

US-09-761-640-4 (1-471) x BE873337 (1-897)

QY 53 AspAsnAspAspAlaAlaGlnAlaSerSerGluProThrGluLysAlaProSerGluGln 72
 DB 2 GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 61
 QY 73 GluLeuHisGlyAspGlnThrAspPheGlyGlnGlySerGlnSerProGlnLysGlnGlu 92
 DB 62 GAGCTCCAGCGGGGACAGACACTTCGGGCAAGGATCCAGAGTCCCAAGAGCAGAGG 121
 QY 93 GlnGlnArgGlnHisLeuHisLeuMetValGlnLeuLeuArgProGlnAspAlaArg 112
 DB 122 GAGCAGAGGACAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 181
 QY 113 LeuAlaAlaGlnLeuGlnAlaProArgProProArgLeuArgTyrLeuLeuValValSer 132
 DB 182 CTGGCAGCCAGCTGAGAGCAGCCCGGCTCCCGGCTCCGCTGCTGCTGCTGCTGCTGCT 241
 QY 133 ThrArgGluGlyGlyGlyLeuSerGlnAspGluThrValLeuLeuGlyValAlaSphePro 152
 DB 242 ACACGAGAGGAGGAGGATGAGCAGGATGAGAGGCTCTCTGCGGCTGATTTCCCT 301
 QY 153 AspSerSerProSerCysThrLeuGlyLeuValLeuProLeuThrSerAspThrGln 172
 DB 302 GACAGCAGCTCCCGCAGCTGACACCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 361
 QY 173 ValTyrLeuAspGlyAspGlyLysPheSerValThrSerGlyGlyGlnSerArgLeuPhe 192
 DB 362 GTGTACTTAGATGAGAGAGCGGGCTTACCGCTGAGCTGTGGGCAAGCGGATCTTC 421
 QY 193 LysProLysSerThrLeuThrMetTrpAlaThrLeuGlnValLeuHisGlnAlaCysGlu 212
 DB 422 AAGCCATCTCCATCCAGACCATGTGGGCGCACATCCAGATGATTCACCAACATGTGAG 481

QY 213 A1a1a1a1euglySerGlyLeuValProGlyGlySerAlaLeuThrTrpAlaSerHisTyr 232
 |||
 DB 482 CGAGCTCAGGAGCGGCTTGACCGGGTGGCAGTCCCTCAGCTGGGCGACCCACTAC 541
 QY 233 G1ng1ua1rLeuAsnSerG1ng1uG1nSerG1yLeuAsnG1uTrpPhrAlaMetAlaAspLeu 252
 |||
 DB 542 CAGGAGAGACTGAACTCCAGAACAGACTGCTC-AATGAGTGGACGGT-ATGGCCGACTGG 599
 QY 253 G1uSerLeuArpProProSerAlaG1uProG1yG1ySerSerG1uG1nG1uG1nMetG1u 272
 |||
 DB 600 ---AGCTCGGGGGCTCCAGCGCCGAGCT-CGGGGGTCTCCCAACAAGAGCCGATGGG- 654
 QY 273 G1na1a1leArq1aG1uLeuTrp1ySvalLeuAspVal 285
 |||
 DB 655 GCAGCGATCCGTGCTGAGCTGGGACGTGTGGATGTG 693
 RESULT 14
 LOCUS BG753072 663 bp mRNA linear EST 15-MAY-2001
 DEFINITION 602732341P1 NIH_MGC_43 Homo sapiens CDNA clone IMAGE:4875562 5',
 mRNA sequence.
 ACCESSION BG753072
 VERSION BG753072.1 GI:14063725
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 663)
 NIH-MGC http://mgs.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1CM1757 row: b column: 11
 High quality sequence stop: 663.
 Location/Qualifiers
 FEATURES
 source
 1. 663
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4875562"
 /clone_lib="NIH_MGC_43"
 /tissue_type="normal pigmented retinal epithelium"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: eye; Vector: pOT87; Site: 1: XhoI; Site: 2:
 EcoRI; CDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGCACGAG(G). Library constructed by Ling Hong
 in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-CDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library.!"
 BASE COUNT 124 a 205 c 231 g 103 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 2.93e-95 Length: 663
 Score: 1009.00 Matches: 197
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 41.42% Indels: 0
 DB: 12 Gaps: 0
 US-09-761-640-4 (1-471) x BG753072 (1-663)
 QY 1 Meta1a1eua1ThrValSerArGserProProG1ySerG1yAlaSerThrProValGly 20

DB 71 ATGGCCCTGTCACAGTGAAGCCGTCGCCCGGGAGGCGGGCTCCAGCCCGGGGG 130
 QY 21 ProTrAspG1a1a1a1a1G1na1rArGserArG1yLeuG1na1rArG1nSerPha1a1a1 40
 |||
 DB 131 CCTGGGACCAAGCGGTCCAGAGAGAGTGCATCCACAGAGAGAGACTTTGGGGTG 190
 QY 41 LeuArG1yAlaValLeuG1yLeuG1nAspG1yG1yAspAsnAspAspAla1a1a1a1a 60
 |||
 DB 191 CTCCTGGGGGTCTCTCTGGAGTGCAGATGAGGGGAGCAATGATGATGCAGAGAGCC 250
 QY 61 SerSerG1uArProThrG1uLysAlaProSerG1uG1uG1uG1nHisG1ySpg1nThrAsp 80
 |||
 DB 251 AGTTCTGACCAACAAGAGAGGCCCGAGTGAAGAGAGACTCCACGGGAGACAGACAGAC 310
 QY 81 PheG1yG1nG1ySerG1nSerProG1nLysG1nG1uG1nArG1nHisLeuHisLeu 100
 |||
 DB 311 TTGGGCAAGATCCACAGAGTCCCCCAAGACAGAGAGAGACAGACACTGCACCTC 370
 QY 101 MetValG1nLeuLeuArpProG1nAspAsp1eArG1yLeuAla1a1a1a1a1a1a1aPro 120
 |||
 DB 371 ATGTACAGCTGCTGAGCGCGCAGATGACATCCGCTGGACCCAGCTGGAGGCACCC 430
 QY 121 ArgProProArG1yLeuArG1yTrLeuLeuVala1SerThrArG1uG1yG1uG1yLeuSer 140
 |||
 DB 431 CGGCTCCCGCGCTCCGCTACGCTGTAATTCTACACGAAAGAGAGAGAGTCTGAC 490
 QY 141 G1naSpG1uThrValLeuLeuG1yValaSpPheProAspSerSerProSerG1yThr 160
 |||
 DB 491 CAGGATGACAGGCTCTCTCGGCGCTGGATTTCCCTGACAGCAGCTCCCGACCTCAC 550
 QY 161 LeuG1yLeuValLeuProLeuTrpSerAspTrpG1nValTyTrLeuAspG1yAspG1yG1y 180
 |||
 DB 551 CTGGGCTGCTGTGGCCCTCTGAGTGAACACCCAGGTACTTACATGAGAGCGGGGCG 610
 QY 181 PheSerValThrSerG1yG1nSerArG1yLeuPhyPro1leSer1e 197
 |||
 DB 611 TTCAGGCTGACGCTGTGGGCAAGCCGATCTTCAACCCATCTCATC 661
 RESULT 15
 LOCUS BE907514 716 bp mRNA linear EST 20-OCT-2000
 DEFINITION 601497473P1 NIH_MGC_70 Homo sapiens CDNA clone IMAGE:3899418 5',
 mRNA sequence.
 ACCESSION BE907514
 VERSION BE907514.1 GI:10401153
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 716)
 NIH-MGC http://mgs.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1AM9697 row: a column: 19
 High quality sequence stop: 683.
 Location/Qualifiers
 FEATURES
 source
 1. 716
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3899418"
 /clone_lib="NIH_MGC_70"
 /tissue_type="epithelioid carcinoma"

```

/lab_host="DH10B (phage-resistant)"
/Note="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.1 kb. Library constructed by Life
Technologies."
BASE COUNT      138 a      213 c      251 g      114 t
ORIGIN

```

```

Alignment Scores:
Pred. No.:      3.7e-93      Length:      716
Score:          989.50      Matches:      207
Percent Similarity: 97.64%      Conservative: 0
Best Local Similarity: 97.64%      Mismatches: 5
Query Match:      40.62%      Indels:      4
DB:              12      Gaps:      0

```

US-09-761-640-4 (1-471) x BE907514 (1-716)

```

OY      1  MetAlaLeuValThrValSerArgSerProGlySerGlyAlaSerThrProValGly 20
      |||
DB      77  ATGGCCCTGTCACAGTAGCGCGTTCGCCCCGGGAGCGGGCCCTCCACGCCCGTGGGG 136
OY      21  ProTTPASPGLnAlaValGlnArgArgSerArgLeuGlnArgGlnSerPheAlaVal 40
      |||
DB      137  CCTGGGACACGGCGGTCAGGAGGAGTCCAGCTCCACGAGGCGAGCTTTGGCGTG 196
OY      41  LeuArgGlyAlaValLeuGlyLeuGlnAspGlyGlyAspAsnAspAlaAlaGluAla 60
      *
DB      197  CTCCTGGGGGCTGCTCTGGGATGTCAGGATGAGGGGACATGATGATGATCAGCAGAGGCC 256
OY      61  SerSerGluProThrGluLysAlaProSerGluGluGluLeuHisGlyAspGlnThrAsp 80
      |||
DB      257  AGTTCAGACCAACAGAGAGAGGCCCGAGTGAAGAGAGAGCTCCACGGGAGCAGACAGAC 316
OY      81  PheGlyGlnGlySerGlnSerProGlnLysGlnGluGlnArgGlnHisLeuHisLeu 100
      |||
DB      317  TTCGGGCAAGATCCAGATCCAGTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 376
OY      101  MetValGlnLeuLeuArgProGlnAspAspIleArgLeuAlaAlaGlnLeuGlnAlaPro 120
      |||
DB      377  ATGTACACACTCTGAGGCGCGAGATGACATCCCGCTGGCAGCCAGCTGAGAGGACACC 436
OY      121  ArgProProArgLeuArgGlyLeuLeuValValSerThrArgGluGlyLeuSer 140
      |||
DB      437  CGGCTCCCGCGGCTCCGCTACCTGCTGAGTTTCTACACGAGAGAGAGAGAGGTTGAGCC 496
OY      141  GlnAspGluThrValLeuLeuGlyValAspPheProAspSerSerSerProSerCysThr 160
      |||
DB      497  CAGGATGAGACGCTCTCTGGGCGTGGATTTCCCTGACAGCAGCTCCCGCAGCTGCACC 556
OY      161  LeuGlyLeuValLeuProLeuThrSerAspThrGlnValThrLeuAspGlyAspGlyGly 180
      |||
DB      557  CTGGGCTGTGCTTTCCTTCGAGTGCACCCAGGTACTTAGATGAGAGCGGGGC 616
OY      181  PheSerValThrSerGlyGlnSerArgIlePheLysProIleSerIleGlnThrMet 200
      |||
DB      617  TTCACGCTGACGCTGAGGGGCAAG-CGGATCTCAACACCATC-TCCATCCAGAC-ATG 673
OY      201  TrpAlaThrLeuGlnValLeuHisGlnAlaCysGln 212
      |||
DB      674  TGGGGACACTCCAGGTATTGGC-CAAGCATGTGAA 708

```

Search completed: April 12, 2003, 10:52:46
Job time : 3880 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame-plus-p2n model

Run on: April 11, 2003, 06:56:50 ; Search time 344 Seconds
(without alignments)
1201.006 Million cell updates/sec

Title: US-09-761-640-4
2436
Sequence: 1 MATVSRSPGSGASTPVG.....PNEGFLROLIYQGLTART 471

Scoring table:
BLOSUM62
Xgapop 10.0, Ygapext 0.5
Fgapop 10.0, Fgapext 0.5
Delop 6.0, Delext 7.0

Searched: 593429 seqs, 438583890 residues

Total number of hits satisfying chosen parameters: 1186858

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n model -DEV=xlh
-O=/cgn2.1/USFO_spool/US09761640/unat_08042003.141440.20481/app.query.fasta.1.647
-DB=Published.Applications_NA -OFMT=fastap -SUFFIX=trnpb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09761640.eccn1.1.99.eccnat.08042003.141440.20481
-NCPU=6 -ICPU=3 -NO_XLPTX -NO_MAP -LARGEOUT -NEG_SCORES=0 -WAIT -LONGLOG
-DEV.TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -DELop=6 -DELEXT=7

Database : Published.Applications_NA:*

1: /cgn2_6/ptodata/2/pubna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubna/PC1_NEM_PUB.seq:*
3: /cgn2_6/ptodata/2/pubna/US06_NEM_PUB.seq:*
4: /cgn2_6/ptodata/2/pubna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubna/US07_NEM_PUB.seq:*
6: /cgn2_6/ptodata/2/pubna/PC1US_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubna/US08_NEM_PUB.seq:*
8: /cgn2_6/ptodata/2/pubna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubna/US09_NEM_PUB.seq:*
10: /cgn2_6/ptodata/2/pubna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubna/US10_NEM_PUB.seq:*
12: /cgn2_6/ptodata/2/pubna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubna/US60_NEM_PUB.seq:*
14: /cgn2_6/ptodata/2/pubna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2436	100.0	1416	9	US-09-955-732-20	Sequence 20, Appl
2	2436	100.0	2704	10	US-09-761-640-1	Sequence 1, Appl
3	2436	100.0	2704	10	US-09-761-640-7	Sequence 7, Appl
4	2426	99.6	1980	9	US-09-955-732-1	Sequence 1, Appl

Result No.	Score	Match	Length	DB	ID	Description
5	2426	99.6	2852	10	US-09-761-640-2	Sequence 2, Appl
6	2061.5	84.6	2540	10	US-09-761-640-3	Sequence 3, Appl
7	1045	42.9	1711	10	US-09-775-925-5	Sequence 5, Appl
8	980	40.2	1949	10	US-09-775-925-1	Sequence 1, Appl
9	589	24.2	1052	10	US-09-775-925-7	Sequence 7, Appl
10	539	22.1	969	10	US-09-808-701-12	Sequence 12, Appl
11	256	10.5	654	10	US-09-815-419-3	Sequence 3, Appl
12	256	10.5	1292	10	US-09-815-419-1	Sequence 1, Appl
13	235.5	9.7	1337	9	US-09-964-899-46	Sequence 46, Appl
14	230	9.4	1398	10	US-09-878-574-1411	Sequence 1411, Ap
15	227.5	9.3	1998	10	US-09-816-494-3	Sequence 3, Appl
16	227.5	9.3	2470	10	US-09-964-824A-555	Sequence 355, Ap
17	227.5	9.3	2470	10	US-09-880-107-3310	Sequence 3310, Ap
18	227.5	9.3	3496	10	US-09-964-277-1	Sequence 1, Appl
19	227.5	9.3	3544	10	US-09-816-494-1	Sequence 1, Appl
20	227	9.3	2436	10	US-09-917-800A-485	Sequence 485, Ap
21	226.5	9.3	2139	9	US-09-865-993-3	Sequence 3, Appl
22	226.5	9.3	2479	9	US-09-865-993-10	Sequence 10, Appl
23	222.5	9.1	705	10	US-09-963-204-3	Sequence 3, Appl
24	222.5	9.1	867	10	US-09-963-204-13	Sequence 13, Appl
25	222.5	9.1	1160	10	US-09-963-204-1	Sequence 1, Appl
26	216.5	8.9	3059	10	US-09-770-595A-2	Sequence 2, Appl
27	216.5	8.9	6356	10	US-09-770-595A-1	Sequence 1, Appl
28	215.5	8.8	2000	10	US-09-969-708-276	Sequence 276, Ap
29	215.5	8.8	2000	10	US-09-919-497-10	Sequence 10, Appl
30	215.5	8.8	2000	10	US-09-880-107-3774	Sequence 3774, Ap
31	215.5	8.8	2303	10	US-09-880-107-3849	Sequence 3849, Ap
32	215	8.8	2064	9	US-09-736-457-825	Sequence 825, App
33	215	8.8	2064	9	US-09-902-941-825	Sequence 825, App
34	215	8.8	2064	9	US-09-849-626-825	Sequence 825, App
35	215	8.8	2064	9	US-10-017-754-825	Sequence 825, App
36	215	8.8	2109	9	US-09-736-457-826	Sequence 826, App
37	215	8.8	2109	9	US-09-902-941-826	Sequence 826, App
38	215	8.8	2109	9	US-09-849-626-826	Sequence 826, App
39	215	8.8	2109	9	US-10-017-754-826	Sequence 826, App
40	215	8.8	2447	9	US-10-071-754-826	Sequence 29, Appl
41	214	8.8	1444	10	US-09-738-885-1	Sequence 1, Appl
42	214	8.8	1691	9	US-09-860-670-19	Sequence 19, Appl
43	209.5	8.6	2453	12	US-10-005-858-1	Sequence 1, Appl
44	209.5	8.6	3332	10	US-09-964-277-20	Sequence 20, Appl
45	207	8.5	1238	9	US-09-736-457-803	Sequence 803, App

ALIGNMENTS

RESULT 1
US-09-955-732-20
; Sequence 20, Application US/09955732
; Publication No. US20020182203A1
; GENERAL INFORMATION:
; APPLICANT: Lucbe, Ralf M.
; TITLE OF INVENTION: DSP-15 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.433
; CURRENT APPLICATION NUMBER: US/09/955,732
; CURRENT FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 1416
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-955-732-20

Alignment Scores:

Pred. No.: 2.64e-258 Length: 1416
Score: 2436.00 Matches: 471
Percent Similarity: 100.00% Conservatve: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-761-640-4 (1-471) x US-09-955-732-20 (1-1416)

```

QY      1 MetAlaLeuValThrValSerArgSerProProGlySerGlyAlaSerThrProValGly 20
      1 ATGGCCCTGGGTCACAGAGAGCGGTTCCGCCCGGGGACAGCGCGCTCCACCCCGGTGGG 60
QY      21 ProTrrAspGlnAlaValGlnArgArgSerArgLeuGlnArgArgGlnSerPheAlaVal 40
      61 CCCGGGACACAGGCGGTCCAGCGAAGAGATCCACTCCAGCCAAAGGACGAGCTTGGCGTG 120
QY      41 LeuArgGlyAlaValLeuGlyLeuGlnAspGlyLysAspAsnAspAspAlaGlnAla 60
      121 CTCGGTGGGCTGTCTCTGGGACTCAGAGATGAGGAGGAGCAATGATGACAGCAGAGGCC 180
QY      61 SerSerGluProThrGluLysAlaProSerGluGluGluLeuHisGlyAspGlnThrAsp 80
      181 AGTTCTGAGCAACAGAGAAAGGCCCGAGTGAAGAGAGGCTCCAGGGGAGCAACAGAGAC 240
QY      81 PheGlyGlnGlySerGlnSerProGlnLysGlnGluGlnArgGlnHisLeuHisLeu 100
      241 TTGGGGCAAGGATCCAGAGTCCCGACAGACAGAGAGAGAGAGGACGACCTGCACCTC 300
QY      101 MetValGlnLeuLeuValProGlnAspAspIleArgLeuAlaGlnLeuGlnAlaPro 120
      301 ATGGTACAGCTGCTGAGGCGGACAGATGACATCCGCTGGAGCCAGCTGGAGGACACC 360
QY      121 ArgProArgLeuArgTyrLeuLeuValValSerThrArgGluGlyGluGlyLeuSer 140
      361 CGGCTCCCGGGCTCCGCTACCTGCTGTGATTCTTACAGAGAGAGAGAGAGAGCTGACG 420
QY      141 GlnAspGluThrValLeuLeuGlyValAspPheProAspSerSerSerProSerCysThr 160
      421 CAGATTAGACGGTCTCTGGGCGTGAATTTCCCTACAGCAGCTCCCGCAGCTGCACC 480
QY      161 LeuGlyLeuValLeuProLeuTrrPheSerAspThrGlnValTyrLeuAspGlyAspGly 180
      481 CTGGGCTGTCTTGGCCCTCGAGTGAACACCCAGSTGTACTTGAATGAGAGAGGGGGC 540
QY      181 PheSerValThrSerGlyGlyGlnSerArgIlePheLysProIleSerIleGlnThrMet 200
      541 TTCAGCGTACGCTCTGTGGGCAAGCGGATCTTCAAGCCCATCTCCATCCAGACCATG 600
QY      201 TrpAlaThrLeuGlnValLeuHisGlnAlaCysGlnAlaAlaLeuGlySerGlyLeuVal 220
      601 TGGGCCACACTCCAGGATATTGACCAAGCATGTGAGGACACTCTAGGACACGGGCTTGT 660
QY      221 ProGlyGlySerAlaLeuThrTrpAlaSerHisTyrGlnGluArgLeuAsnSerGluGln 240
      661 CCGGTTGGGAGTGCCCTCACCCTGGGCCAGCCTACACAGAGAGAGACTGGAAGCTCCGAG 720
QY      241 SerCysLeuAsnGluTrrPheAlaMetAlaAspLeuGlnSerLeuArgProProSerAla 260
      721 AGCTGCTCATATGATGAGGAGGCTATGGCCGACTGTGAGTCTGCGGCTCCAGCGCC 780
QY      261 GluProGlyLysSerSerGluGlnGluGlnMetGluGlnAlaIleArgAlaGluLeuTrr 280
      781 GAGCCCTGGGCGGCTCTCAGACAGAGACAGATGAGAGGCGATCCCTGTGAGCTGTGG 840
QY      281 LysValLeuAspValSerAspLeuGluSerValThrSerLysGluIleArgGlnAlaLeu 300
      841 AAGTGTGATGTCAGTGACTGAGAGAGTGTACTTCCAAAGAGATCCCGCAGGCTCTG 900
QY      301 GluLeuArgLeuGlyLeuProLeuGlnGlnTyrArgAspPheIleAspAsnGlnMetLeu 320
      901 GAGCTCGGCTGGGCTCCCGCTCCAGACAGTACCGTACTTCATCGACCAACAGATGCTG 960
QY      321 LeuLeuValAlaGlnArgAspArgAlaSerArgIlePheProHisLeuTyrLeuGlySer 340
      961 CTGCTGTGTGACAGCGGAGCGAGGCTCCCGCATCTTCCCGACCTTACCTGCGGCTCA 1020
QY      341 GluTrrAsnAlaAlaAsnLeuGluGluLeuGlnArgAsnArgValThrHisIleLeuAsn 360
      1021 GAGTGAACCGCACCAACCTGGAGAGAGCTGCAAGAGACAGGGGTCAACCCACATCTTGAAC 1080

```

```

QY      361 MetAlaArgGluIleAspAsnPheTyrProGluArgPheThrTrrHisAsnValArgLeu 380
      1081 ATGGCCCGGAGAGATTGACAACTTACCTTACCGGCTTCACTTCAACCAATGTGGCGCTC 1140
QY      381 TrpAspGluSerAlaGlnLeuLeuProHisTrrPlysGluThrHisArgPheIleGlu 400
      1141 TGGGATGAGAGTGGGCCCAAGCTCTCGCACTGGAAGAGAGACGACCGCTTCAATTGAG 1200
QY      401 AlaAlaArgAlaGlnGlyThrHisValLeuValHisCysLysMetGlyValSerArgSer 420
      1201 GCTGCAAGAGACAGGGCCACCCAGTGTCTGTCCACTGCAAGATGAGGCGTCAAGCGCTCA 1260
QY      421 AlaAlaThrValLeuAlaTyrAlaMetLysGlnTrrGlySerLeuGlnGlnAlaLeu 440
      1261 GCGGCCACAGTGTGGCTTATGCCATGACAGACAGTACGAATGCAACCTCGAGAGAGCCCTG 1320
QY      441 ArgHisValGlnGluLeuArgProIleAlaArgProAsnProGlyPheLeuArgGlnLeu 460
      1321 CGCCAGCTGACAGAGCTCCGCGCCATCGCCGCCCAACCTGGCTTCTGCGCAGCTG 1380
QY      461 GlnIleTyrGlnGlyIleLeuThrAlaArgThr 471
      1381 CAGATCTACCAAGGCGATCTGACGGCCAGAAC 1413

```

RESULT 2

```

US-09-761-640-1
; Sequence 1, Application US/09761640
; Patent No. US20020137042A1
;
; GENERAL INFORMATION:
; APPLICANT: MET, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHATASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHATASE PROTEINS,
; FILE OF INVENTION: AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/09/761,640
; CURRENT FILING DATE: 2001-01-18
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2704
; TYPE: DNA
; ORGANISM: Human
US-09-761-640-1

```

Alignment Scores:

```

Pred. No.:      6,61e-258      Length:      2704
Score:          2436.00      Matches:      471
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      100.00%      Indels:      0
DB:               Gaps:      0

```

US-09-761-640-4 (1-471) x US-09-761-640-1 (1-2704)

```

QY      1 MetAlaLeuValThrValSerArgSerProProGlySerGlyAlaSerThrProValGly 20
      1 ATGGCCCTGGTCAAGAGAGAGCGGTTCCGCCCGGGGACAGCGCGCTCCACCCCGGTGGG 153
QY      21 ProTrrAspGlnAlaValGlnArgArgSerArgLeuGlnArgArgGlnSerPheAlaVal 40
      61 CCCGGGACACAGGCGGTCCAGCGAAGAGATCCACTCCAGCCAAAGGACGAGCTTGGCGTG 213
QY      41 LeuArgGlyAlaValLeuGlyLeuGlnAspGlyLysAspAsnAspAspAlaGlnAla 60
      121 CTCGGTGGGCTGTCTCTGGGACTCAGAGATGAGGAGGAGCAATGATGACAGCAGAGGCC 273
QY      61 SerSerGluProThrGluLysAlaProSerGluGluGluLeuHisGlyAspGlnThrAsp 80
      181 AGTTCTGAGCAACAGAGAAAGGCCCGAGTGAAGAGAGGCTCCAGGGGAGCAACAGAGAC 333
QY      81 PheGlyGlnGlySerGlnSerProGlnLysGlnGluGlnArgGlnHisLeuHisLeu 100
      241 TTGGGGCAAGGATCCAGAGTCCCGACAGACAGAGAGAGAGAGGACGACCTGCACCTC 393

```

OY	101	MethylGluLeuLeuArgProGlnAspPheIleArgLeuAlaAlaGlnLeuGluAlaPro	120
Db	394	ATGGTACAGCTGCTGAGGCGCGAGATGATCAATCCGCTGGACGCCACTGGAGGACACC	453
OY	121	ArgProPheArgLeuArgTyrLeuLeuValIleSerThrArgGluGlyGluGlyLeuSer	140
Db	454	CGGGCTCCCGGGCTCCGCTACCTGCTGGTATGTTTCTACACGAGAAAGAAAGTCTGAGC	513
OY	141	GlnAspGluThrValLeuLeuGlyValAspPheProAspSerSerProSerCysThr	160
Db	514	CAGATGAGACGGCTCCCTGGGCGTGGATTTCCCTGACAGCAGCTCCGCCAGCTGCACC	573
OY	161	LeuGlyLeuValLeuProLeuTyrPheSerPheThrGlnValTyrLeuAspGlyAspGlyLy	180
Db	574	CTTGAGCTGGCTTGGCCCTCTGGAGTGCACCCAGCTGTACTTGATGATGAGACGGGGC	633
OY	181	PheSerValThrSerGlyGlnSerArgIlePheLyProIleSerIleGlnThrMet	200
Db	634	TTACGGCGTACGCTGTGGTGGGCAACCGGATCTTCAACCCATCTCATCAAGCATG	693
OY	201	TrpAlaThrLeuGlnValLeuHisGlnAlaCysGluAlaAlaLeuGlySerGlyLeuVal	220
Db	694	TGGCGCACACTCCAGGTATTGCACCAAGCATGTGAGGCGACTGTAGGACGGGCTTGTA	753
OY	221	ProGlyGlySerAlaLeuThrTrpAlaSerHisTyrGlnGluArgLeuAsnSerGlnGln	240
Db	754	CCGGGTGGCACTGGCCCTCAGCTGGCGGACGCACTACCAAGAGACTGAATCTCGAATG	813
OY	241	SerCysLeuAsnGluTrpThrAlaMetAlaAspLeuGluSerLeuArgProProSerAla	260
Db	814	AGCTGGCCATATGATGTGAGCGGCTATGGCGACCTGATGATCTCTCGGCTCCAGCCG	873
OY	261	GluProGlyGlySerSerGluGlnGluGlnMetGluGlnAlaIleArgAlaGlnLeuTrp	280
Db	874	GAGCTTGCGGGGCTCTCAGACAGGAGCAGATGAGACAGGCTCGGTGAGCTGTGG	933
OY	281	LysValLeuAspValSerAspLeuGluSerValThrSerGlyGluIleArgGlnAlaLeu	300
Db	934	AAAGTTGGATGTCACTGATGACCTGGAGAGTGCACCTCCAAAGATCCCGCAGCTCTG	993
OY	301	GluLeuArgLeuGlyLeuProLeuGlnGlnTyrArgAspPheIleAspAsnGlnMetLeu	320
Db	994	GAGCTGGCCCTGGGGCTCCCTCCAGCAGTACCGGATCTTATGCAACCAAGTGTGTG	1053
OY	321	LeuLeuValAlaGlnArgAspArgAlaSerArgIlePheProHisLeuTyrLeuGlySer	340
Db	1054	CTGCTGGTGGCACGCGGGACCCAGCTCCCGCATTTTCCCACTCTACTTGGGCTCA	1113
OY	341	GluTrpAsnAlaAlaAsnLeuGlnGluLeuGlnArgAsnArgValThrHisIleLeuAsn	360
Db	1114	GAGTGAACCGACGCAAACTCGGAGAGCTGACAGAGAAACAGGTGCACCATCTTGAAC	1173
OY	361	MetIleArgGluIleAspAsnPheTyrProGluArgPheThrThrHisAsnValArgLeu	380
Db	1174	ATGGCCGGGAGATTGCACAACTTCTACCTGAGCGCTTAACTACACAAATGCGGCTC	1233
OY	381	TrpAspGluSerAlaGlnLeuLeuProHisTyrPheGlyGluThrHisArgPheIleGlu	400
Db	1234	TGGGATGAGGATGGCGCCAGCTGTGGCCGCACTGGAAGAACGACGACGGCTTCATTG	1293
OY	401	AlaAlaArgAlaGlnGlyThrHisValLeuValHisCysLyseMetGlyValSerArgSer	420
Db	1294	GCTGCAAGACACGGGCGACCCAGCTGGTGTCACTGCAAAATGGCGCTCAGCCGTCA	1353
OY	421	AlaAlaThrValLeuAlaTyrAlaMetLysGlnTyrGlyCysSerLeuGlnGlnAlaLeu	440
Db	1354	GCGGCAAGTGTGGCTGTATGGCATGAAGACAGTACAGATGACAGCTGGAGAGGCGCTG	1413
OY	441	ArgHisValGlnGluLeuArgProIleAlaArgProAsnProGlyPheLeuArgGlnLeu	460
Db	1414	CGCCACGTGACAGAGCTCCGGCCATCGCCCGCCCAACCTGTGGTCTCTGGCCAGCTG	1473

QY	461	GlnIleTyrgIngllyIleleuThrAlaArgthr	471
DB	1474	CAGATTCACAGGGCATCTGACGGCCAGAAC	1506
RESULT 3			
	US-09-761-640-7		
	: Sequence 71, Application US/09761640		
	: Patent No. US20020137042A1		
	: GENERAL INFORMATION:		
	: APPLICANT: WEI, Ming-Hui et al		
	: TITLE OF INVENTION: ISOLATED HUMAN PHOSPHATASE PROTEINS,		
	: TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHATASE PROTEINS		
	: TITLE OF INVENTION: AND USES THEREOF		
	: FILE REFERENCE: CLO00964-CIP		
	: CURRENT APPLICATION NUMBER: US/09/761,640		
	: CURRENT FILING DATE: 2001-01-18		
	: NUMBER OF SEQ ID NOS: 10		
	: SOFTWARE: FASTSEQ for Windows Version 4.0		
	: SEQ ID NO 7		
	: LENGTH: 2704		
	: TYPE: DNA		
	: ORGANISM: Human		
	US-09-761-640-7		
Alignment Scores:			
	Pred. No.:	6.61e-258	Length: 2704
	Score:	2436.00	Matches: 471
	Percent Similarity:	100.00%	Conservative: 0
	Best Local Similarity:	100.00%	Mismatches: 0
	Query Match:	100.00%	Indels: 0
	DB:	10	Gaps: 0
US-09-761-640-4 (1-471) x US-09-761-640-7 (1-2704)			
QY	1	MetalAlaValThrValSerArgSerProProGlySerGlyAlaSerThrProValGly	20
DB	94	ATGGCCCTGGGTGCACAGTAGCCGTTCCGCCCCGGCAGCCGGCGCTCCACCCCGTGGG	153
QY	21	ProTPARPGInAlaValAlaGlnArgArgSerArgLeuGlnArgArgGlnSerPheAlaVal	40
DB	154	CCCTGGGACACGCGGGTCCACGAGAGGTGACTCCACGACAGGACAGCTTTGGCGTG	213
QY	41	LeuArgGlyAlaValAlaLeuGlyLeuGlnAspGlyGlyAspAsnAspAspAlaAlaGlyAla	60
DB	214	CTCCGTGGGGCTGTCTCTGGACTGCAGATGGAGGGCACAATGATGATGCAGAGAGGCC	273
QY	61	SerSerGlnProThrGlnLysAlaProSerGlnGlnGlnLysGlnAspGlnThrAsp	80
DB	274	ACTTGTGAGCCACAGAGAGAGGCCGCCGAGTAGAGGAGGAGCTCCACGGGGACACAGAC	333
QY	81	PheGlyGlnGlySerGlnSerProGlnLysGlnGlnGlnGlnGlnArgGlnHisLeu	100
DB	334	TTCCGGGCAAGATCCCAAGATGCCAGAAAGCAGAGGAGGAGGACAGACACTGCACCTC	393
QY	101	MetValGlnLeuLeuLeuArgProGlnAspAspIleArgIleAlaGlnLeuGlnAlaPro	120
DB	394	ATGGTACAGCGTCTGAGGCCCGAGGATACATCCGCCCTGGAGGCCACGTGGAGGACCC	453
QY	121	ArgProProArgLeuArgTyrLeuLeuValAlaSerThrArgGlnGlyGlyLeuSer	140
DB	454	CGGGCTCCCGGGCTCCGTACTCTGTGTAGTTTCTACACAGAGAGGAGAGGTGTGAGC	513
QY	141	GlnAspGlnThrValLeuLeuGlyValAlaAspPheProAspSerSerProSerCysThr	160
DB	514	CAGGATGAGACGTCCTCTCTGGCGCTGATTTCCCTGCACACACTCCGCCACGCTGCACC	573
QY	161	LeuGlyLeuValLeuProLeuThrSerSerPheGlnValTyrLeuAspGlyAspGlyGly	180
DB	574	CTGGGCTGGGTCTTCCCTCTGAGTAGACACCCAGGAGTACTAGAGGAGGAGGGGGGC	633
QY	181	PheSerValThrSerGlyGlyGlnSerArgIlePheIysProIleSerIleGlnThrMet	200
DB	634	TTTCACGTCGAGCTCTGGTGGGCAAGCGCGAATCTTCAAGCCCATCTGCATCCAGACCATG	693

```
QY 201 TTPAlaThrLeuGlnValLeuHisGlnAlaCysGluAlaAlaLeuGlySerGlyLeuVal 220
    |||
Db 694 TGGGCCACACCTCAGGATATTGCACCAAGCATGTGAGGACAGCTCTAGGACGGCCCTTGTA 753
QY 221 PROGlyGlySerAlaLeuThrTTPAlaSerHisTyrGlnGluArgLeuAsnSerGluGln 240
    |||
Db 754 CCGGGTGGCAGTGCCTCCCTCAGCTGGGCCACCATACAGAGAGACTGAACCTCCGAACAG 813
QY 241 SerCysLeuAsnGluThrPThrAlaMetAlaAspLeuGluSerLeuArgProProSerAla 260
    |||
Db 814 AGCTGCTCAATATGATGAGCGGCTATGGCCGACCTGAGTCTTGCGGCTCCAGAGGCC 873
QY 261 GluProGlyGlySerSerGluGlnGlnGlnMetGluGlnAlaLeuArgAlaLeuTTP 280
    |||
Db 874 GAGCTGGCGGGTCTCTCAGAACAGAGAGATGAGCAGGCGATCCGTGTGAGCTGTGG 933
QY 281 LysValLeuAspValSerAspLeuGluSerValThrSerLysGluLeuArgGlnAlaLeu 300
    |||
Db 934 AAAGTGTGGATGTCAGTGACCTGGAGAGTGTCACTCCAAAGAGATCCGCCAGGCTCTG 993
QY 301 GluLeuArgLeuGlyLeuProLeuGlnGlnTyrArgAspPheThrLeuAsnGlnMetLeu 320
    |||
Db 994 GAGCTCGGCTGGGGCTCCCTCCAGCAGTACCTGACTTCAATCCAAACAGATGCTG 1053
QY 321 LeuLeuValAlaGlnArgAspArgAlaSerArgLLePheProHisLeuTyrLeuGlySer 340
    |||
Db 1054 CTGCTGTGGCAGACAGGGGACCGAGCTCCGCGATCTTCCCGCAGCTCTACCTGAGCTCA 1113
QY 341 GluTTPAsnAlaAlaAsnLeuGluGlnGlnArgAsnArgValThrHisLeuAsn 360
    |||
Db 1114 GAGTGAAGCAGCAAACTCGGAGAGCTGCAGAGAGAACAGGGTCAACCATCTTGAAAC 1173
QY 361 MetAlaArgGluThrLeuAspAsnPheTyrProGluArgPheThrHisAsnValArgLeu 380
    |||
Db 1174 ATGGCCCGGAGATTACAACTTCTACCTGAGCGCTTACCTACCAATGCGCCCTC 1233
QY 381 TTPAspGluGluSerAlaGlnLeuLeuProHisTyrLysGluThrHisArgPheLeuGlu 400
    |||
Db 1234 TGGGATGAGAGAGTCGCGCCAGCTGTCGCGCAGTGAAGAGAGACGACCGCTTCATTGAG 1293
QY 401 AlaAlaArgAlaGlnGlyThrHisValLeuValHisCysLysMetLysValSerArgSer 420
    |||
Db 1294 GCTGCAGAGCAGAGGGCACCACGCTGTGTCTCACTGCAGAGATGGCGGTACAGCCCTCA 1353
QY 421 AlaAlaThrValLeuAlaTyrAlaMetLysGlnTyrGluCysSerLeuGlnGlnAlaLeu 440
    |||
Db 1354 GCGGCCACAGTGTGCTGCTATGCCATGAAAGCATGACGATGCGCTGGAGAGCGCCCTG 1413
QY 441 ArgHisValGlnGluLeuArgProHisAlaArgProAsnProGlyPheLeuArgGlnLeu 460
    |||
Db 1414 CCGCAGGTGACAGAGCTCGCGCCATCGCCCGCCCAACCCCTGCTTCCGCGCCAGCTG 1473
QY 461 GlnLLeTyrGlnGlyLLeuThrAlaArgThr 471
    |||
Db 1474 CAGATCTACAGGGCATCTGACGGCCAGAAC 1506

RESULT 4
US-09-955-732-1
; Sequence 1, Application US/09955732
; Publication No. US20020182203A1
; GENERAL INFORMATION:
; APPLICANT: Luche, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-15 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.433
; CURRENT APPLICATION NUMBER: US/09/955,732
; CURRENT FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1980
; TYPE: DNA
```

```
; ORGANISM: Homo sapiens
US-09-955-732-1

Alignment Scores: 5,34e-257 Length: 1980
Pred. No.: 2426.00 Matches: 469
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 99.59% Indels: 0
Query Match: 99.59% Gaps: 0

US-09-761-640-4 (1-471) x US-09-955-732-1 (1-1980)

QY 1 MetAlaLeuValThrValSerArgSerProProGlySerGlyAlaSerThrProValGly 20
    |||
Db 1 ATGGCCCTGTGTCACATGAGCGGTTCGCCCGGGGAGGCGGCTCCACAGCCCGTGGGG 60
QY 21 ProTTPAspGlnAlaValGlnArgArgSerArgLeuGlnArgArgGlnSerPheAlaVal 40
    |||
Db 61 CCTGGGACACGCGCTCCAGCAAGAGACTGCATCCACGCAAGCCAGAGCTTTGCGGTG 120
QY 41 LeuArgGlyAlaValLeuGluGlnGlnAspGlyLysAspAsnAspAlaAlaGluAla 60
    |||
Db 121 CTCCTGGGGCTGTCTGGAGCTGACAGATGGAGGGACAAATGATATGCAGAGAGGCC 180
QY 61 SerSerGluProThrGluLysAlaProSerGluGluGluLeuHisGlyLyspGlnThrAsp 80
    |||
Db 181 AGTTTGAGCAGCAACAGAAAGGCCCGGAGTGAAGAGAGCTCCACGGGAGCAGACAGAC 240
QY 81 PheGlyGlnGlySerGlnSerProGlnLysGlnGlnArgGlnGlnHisLeu 100
    |||
Db 241 TTGCGGCAAGAGATCCAGAGTCCCGAAGCAGAGAGACAGAGGAGCAGCAGCCTGCACTC 300
QY 101 MetValGluLeuLeuArgProGlnAspAspLLeuArgLeuAlaAlaGlnGlnGlnAlaPro 120
    |||
Db 301 ATGGTACAGCTGTGAGGCGCCGAGATGACATCCGCTGACGCCACCCAGCTGAGGACCC 360
QY 301 ArgProProArgLeuArgTyrLeuLeuValValSerThrArgGluGlyLysLeuSer 140
    |||
Db 361 CGGCTCCCGGGCTCGGTACGCTGCTGATGTTTCTACACGAGAAAGAGAGAGTGTGACC 420
QY 141 GlnAspGluThrValLeuLeuGlyValAspPheProAspSerSerSerProSerCysThr 160
    |||
Db 421 CAGGATGAACAGGTCCTCGGGCGTGGAATTCCTCCGACAGACGCCCCCGAGCTGCACC 480
QY 161 LeuGlyLeuValLeuProLeuThrPThrAspThrGlnValTyrLeuAspGlyLysGly 180
    |||
Db 481 CTGGGCTGTGCTTGGCCCTCTGGAATGACACCCAGGTGTACTTGAATGAGACGGGGCC 540
QY 181 PheSerValThrSerGlyGlnSerArgLLePheLysProLLeSerLLeGlnThrMet 200
    |||
Db 541 TTCAGGTGACGTCGTGGTGGGCAAGCGGATCTTCAAGCCCATCTCCATCCAGACATG 600
QY 201 TTPAlaThrLeuGlnValLeuHisGlnAlaCysGluAlaAlaLeuGlySerGlyLeuVal 220
    |||
Db 601 TGGGCCACACTCAGGATATTGCACCAAGCATGAGAGCAGCTCTAGCGAGCGGCTTGTA 660
QY 221 PROGlyGlySerAlaLeuThrTTPAlaSerHisTyrGlnGluArgLeuAsnSerGluGln 240
    |||
Db 661 CCGGGTGGCAGTGCCTCCTCAGCTGGGCCACCACTACCAAGAGAGACTGAATCCGAGACG 720
QY 241 SerCysLeuAsnGluThrPThrAlaMetAlaAspLeuGluSerLeuArgProProSerAla 260
    |||
Db 721 AGTGTGCTCAATGATGAGGAGGGGTATGGCGAGACTGATGATCTCTGGGCTCCAGAGCC 780
QY 261 GluProGlyGlySerSerGluGlnGlnGlnMetGluGlnAlaLeuArgAlaLeuTTP 280
    |||
Db 781 GAGCTGGCGGGTCTCTCAGAACAGAGAGATGAGAGAGCAGGCGATCCGTGTGAGCTGTGG 840
QY 281 LysValLeuAspValSerAspLeuGluSerValThrSerLysGluLeuArgGlnAlaLeu 300
    |||
Db 841 AAAGTGTGATGTCAGTGACCTGGAGAGTGTCACTCCAAAGAGATCCGCGCAGGCTCTG 900
```


US-09-761-640-4 (1-471) x US-09-775-925-1 (1-1949)

```
OY 142 AspGluThrValLeuLeuGlyAlaSpheProAspSerSerProSerGlyThrLeu 161
Db 32 GAGGAGAAATATCTCTGAGTGAATCTTCCAGTAAAGAAATGAAAGGACCAAT 91
OY 162 GlyLeuValLeuProLeuThrSerAspThrGlnValTyrLeuAspGlyAspGlyPhe 181
Db 92 GGGATGGTCTCCGACTGCGAGGACACAGAAATCCACCTTGATGAGATGTGGTTC 151
OY 182 SerValThrSerGlyGlnSerArgIlePheLysProIleSerIleGlnThrMetIrp 201
Db 152 AGCGTAGAGACAGACAGACATGACATATTTAAAGCCGTGTCTGTCCAGGCCATGTG 211
OY 202 AlaThrLeuGlnValLeuHisGlnAlaGlyGluAlaLeuGlySerIleValLeuValPro 221
Db 212 TCTGCCCTGACAGTCTTCACAGGCTTCGAGTGGCCGAGGACCACTACTTCCCC 271
OY 222 GlyGlySerAlaLeuThrTrpAlaSerHisTyrGlnGluArgLeuAsnSerGlnSer 241
Db 272 GGGGGTAGAGCTCTCATCTGGGCTACCTACTGATGAGAGCTGCATCAGCTCCAGAGAGC 331
OY 242 CysLeuAsnGluTrpThrAlaMetAlaAspLeuGluSerLeuArgPro-----ProSer 259
Db 332 TCCATCAACAGAGTGGACGCGCATGCGAGCTGGAGTTCAGCGCGCCGACTCCCGCGG 391
OY 260 AlaGluProGlyGlySerSerGluGlnGluGlnMetGluGlnAlaIleArgAlaGluLeu 279
Db 392 CTAATTGGGACACAGCCCACTGAAGGGAAGGACCGAGCGCTCATCAAGCCCAAGCTTC 451
OY 280 TrpLysValLeuAspValSerAspLeuGluSerValThrSerLysGluIleArgGlnAla 299
Db 452 CCAAGCATCATGATGAGCCAGCATGCAAAATGTCACTTCCAAAGATGCTGTAATGAA 511
OY 300 LeuGluLeuArgLeuGlyLeuProLeuGlnGlnIleTyrArgAspPheIleAspAsnGlnMet 319
Db 512 TTAGGAAACACATGATGATTGTACTTGAAGAACTCAAGAAATTATATACATATGAGAG 571
OY 320 LeuLeuLeuValAlaGlnIleArgAspArgAlaSerArgIlePheProHisLeuTyrIleGly 339
Db 572 CTACTTATGTTGGACACAGATGACAAAGCCCTCCCTTATCTTGATCATCTTTATCTCGGC 631
OY 340 SerGluTrpAsnAlaAlaAsnLeuGluGlnIleArgAsnArgValThrHisIleLeu 359
Db 632 TCTGAATGGAATGCATCCATCTGAGGAACTGACAGGCTCAGGGGTGATTACATTTTA 691
OY 360 AsnMetAlaArgGluIleAspAsnPheTyrProGluArgPheThrTyrHisAsnValArg 379
Db 692 AATGTTACCGAGAAATCGATATATTTTTCCTGCTTATTTGCATATCATATAACATCGA 751
OY 380 LeuTrpAspGluGluSerAlaGlnLeuLeuProHisTrpLysGluThrHisArgPheIle 399
Db 752 GCTTCAGATGAAGAGACACACAGACCTCCGCCCACTGGAATGAACCGATATCTTTA 811
OY 400 GlnAlaAlaArgAlaGlnGlyThrHisValLeuValHisCysLysMetGlyValSerArg 419
Db 812 AACCAAGCGCAAGAGAACCATTCCTCAAGTCCCTGTCATTTGCAAAATGGCGCTGATCCG 871
OY 420 SerAlaAlaThrValLeuAlaTyrAlaMetLysGlnTyrGluCysSerLeuGluGlnAla 439
Db 872 TGGGGCTCCACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 931
OY 440 LeuArgHisValGlnGluLeuArgProIleAlaArgProAsnProGlyPheLeuArgGln 459
Db 932 TTTAACTATGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 991
OY 460 LeuGlnIleTyrGlnGlyIleLeuThrAla 469
Db 992 CTGTCTGAGTATGAGAGCATCTTGATGCA 1021
```

RESULT 9
US-09-775-925-7
Sequence 7, Application US/09775925

```
; Patent No. US20010049358A1
; GENERAL INFORMATION:
; APPLICANT: Lucbe, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-12 AND DSP-13 DUAL-SPECIFICITY
; TITLE OF INVENTION: PHOSPHATASES
; FILE REFERENCE: 200125.420
; CURRENT APPLICATION NUMBER: US/09/775,925
; CURRENT FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1052
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-775-925-7
```

```
Alignment Scores:
Pred. No.: 1,94e-55 Length: 1052
Score: 589.00 Matches: 107
Percent Similarity: 75.35% Conservative: 55
Best Local Similarity: 49.77% Mismatches: 45
Query Match: 24.18% Indels: 8
DB: 10 Gaps: 1
```

US-09-761-640-4 (1-471) x US-09-775-925-7 (1-1052)

```
OY 255 LeuArgProProSerAlaGluProGlyGlySerSerGluGlnGluMetGluGlnAla 274
Db 190 CTCAGACCT-----ACGAAAGTGAAGAGACAGAGAAAGGCTA 225
OY 275 IleArgAlaGluLeuThrTrpLysValLeuAspValSerAspLeuGluSerValThrSerLys 294
Db 226 ATTTAAACCAAAATTAAGGAGATCATGATGATGATGATGATGATGATGATGATGATGATG 285
OY 295 GlnIleArgGlnAlaLeuGlnIleuArgLeuGlyLeuProLeuGlnGlnIleTyrArgAspPhe 314
Db 286 GAGATTAAGACAGAGATGGAATGCAAAATGCTGCACACTTCGGGAATTCAGGAATTT 345
OY 315 IleAspAsnGlnMetLeuLeuLeuValAlaGlnArgAspArgAlaSerArgIlePhePro 334
Db 346 ATAGACAAATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 405
OY 335 HisLeuTyrLeuGlySerGluTrpAsnAlaAlaAsnLeuGluGlnIleuArgAsnArg 354
Db 406 CATGCTTCTCGGGCTCAGAAATGGAATGCCCTCACTTGAAGAGACTACAGAACGAGG 465
OY 355 ValThrHisIleLeuAsnMetAlaArgGluIleAspAsnPheTyrProGluArgPheThr 374
Db 466 GTACGGTATATCTTGAATGCTACGAGAGATGATGATGATGATGATGATGATGATGATG 525
OY 375 TyrHisAsnValArgLeuTrpAspGluGlnSerAlaGlnLeuLeuProHisTrpLysGln 394
Db 526 TATCATTAACATTCGGGTATATGATGATGATGATGATGATGATGATGATGATGATGATG 585
OY 395 ThrHisArgPheIleGlnAlaAlaArgAlaGlnGlyThrHisValLeuValHisCysLys 414
Db 586 ACTTACAAATTCATCTCTAAACAAAGAAACATGATGATGATGATGATGATGATGATG 645
OY 415 MetGlyValSerArgSerAlaThrValLeuAlaTyrAlaMetLysGlnTyrGluCys 434
Db 646 ATGGGGGTAGTGGCTCACACCTCCACCGTATGCTTATGATGATGATGATGATGATG 705
OY 435 SerLeuGlnGlnAlaLeuArgHisValGlnGluLeuArgProIleAlaArgProAsnPro 454
Db 706 AATCTGACCGAGCCCTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 765
OY 455 GlyPheLeuArgGlnLeuGlnIleTyrGlnGlyIleLeuThrAla 469
Db 766 AGCTTCATGACAACTGGAAGAGATATCAGGGAGATCTTCTGCGCA 810
```

RESULT 10
US-09-808-701-12

```
; Sequence 12, Application US/09808701
; Patent No. US20020146757A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Xue, Aidong J.
; APPLICANT: Ren, Feiyan
; APPLICANT: Wang, Dunrui
; APPLICANT: Chen, Rui-hong
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20020146757A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 790CIP2D
; CURRENT APPLICATION NUMBER: US/09/808,701
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PL_FL_genes Version 2.0
; SEQ ID NO 12
; LENGTH: 969
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (34)..(765)
US-09-808-701-12

Alignment Scores:
Pred. No.: 5,43e-50 Length: 969
Score: 539.00 Matches: 120
Percent Similarity: 59.55% Conservative: 39
Best Local Similarity: 44.94% Mismatches: 70
Query Match: 22.13% Indels: 38
DB: 10 Gaps: 7

US-09-761-640-4 (1-471) x US-09-808-701-12 (1-969)
QY 1 MetAlaLeuValThrValSerArgSerProGlySerGlyAlaSerThrProValGly 20
   |||||||
DB 34 ATGGCCCTGGTGCAGCTGCAGCCGCGCCACGCCCGCCCTCTCGGCGCAGC 93
QY 21 ProTPrAsp-----GlnAlaValGlnArgSerArgLeuGlnArgGln 36
   :|:|:|
DB 94 AACAGCGAGTGTGGAGCTGGCAGCGAAGAAGATCGAAATTAACCTCAGCTTAAGTGA 153
QY 37 SerPheAlaValLeuArgGlyAlaValLeuGlyLeuGlnAspGlyGlyAspAsnAsp 56
   ||||| :|:|:|:|
DB 154 ACCTTTTTCATGTGTAAGCGCGCAGCCCTCTCTTACACAGGGA----- 198
QY 57 AlaAlaGlnAlaSerSerLysProThrGluLysAlaProSerGluGluGluLeuHisGly 76
   ||| |||
DB 199 -----AGCAGCCCTCAA----- 210
QY 77 AspGlnThrAspPheGlyGlnGlySer---GlnSerProGlnLysGlnGlu---GlnGln 94
   ||||| ||| ||| ||| |||
DB 211 -----GGCCAGCGGAGTCTTCAGCAACCCACACAGATCAGGTGATCTG 255
QY 95 ArgGlnHisLeuHisLeuMetValGlnLeuLeuArgProGlnAspSpleArgLeuAla 114
   ||||| :|:|:|:|
DB 256 CCTCAACATCTTCAGGTGATGATCAACTTCTGCGTGCAGAGACAGATCAAGCTGCA 315
QY 115 AlaGlnLeuGlnAlaProArgProProArgLeuArgTyrLeuLeuValAla-----Ser 132
   :|:|:|:|
DB 316 GTGGCGGTGAGAGCGCCCTGGGCGGAGCCGGGTCCGTACATGAGTGAACACAGCC 375
QY 133 ThrArgGluGlyGluLysLeuSerGlnAspGluThrValLeuLeuGlyAlaAspPhePro 152
```

```
DB 376 GGGCCGCGAG-----GACACCGAGGAGAAATATCTGCTGAGTGGACTTTCC 423
   |||:|:|
QY 153 AspSerSerSerProSerCysThrLeuGlyLeuValLeuProLeuTyrSerAspThrGln 172
   ||| |||||
DB 424 AGTAAAGAAAGTAAAGAGCGTGCACACATTGGATGTTCTCCAGTGGAGCGACAGAA 483
QY 173 ValTyrLeuAspGlyAspGlyLysPheSerValThrSerGlyGlyGlnSerArgIlePhe 192
   :|:|:|:|
DB 484 ATCCACCTTGATGAGATGCTGGTTTCACCGTCGACACAGCAAGAAAGATGCACATATT 543
QY 193 LysProIleSerTleGlnThrMetThrPalaThrLeuGlnValLeuHisGlnAlaCysGlu 212
   |||||:|:|:|
DB 544 AAGCTGTGTCTGTCCAGGCGCATGTGTTCTGCCCTGCAGAGTCTTCACAAAGGCTGCGAA 603
QY 213 AlaAlaLeuGlySerGlyLeuValProGlyGlySerAlaLeuThrThrPalaSerHisTyr 232
   |||||
DB 604 GTGGCCCGGAGGCACACACTTCTCCCGGGGGTGTAGCTCTCATCTGGGCTACTTACT 663
QY 233 GlnGluArgLeuAsnSerGlnLysSerCysLeuAsnGluTyrThrAlaMetAlaAspLeu 252
   :|:|:|:|
DB 664 GAGAGCTGCATCAGCTCCGACGACAGCTTCATCAACGATGAGACGCCATGACGACCTG 723
QY 253 GluSerLeuArgProProSer 259
   |||||
DB 724 GAGTCTACGCGGCGGCGACTCC 744

RESULT 11
US-09-815-419-3
; Sequence 3, Application US/09815419
; Patent No. US20020065406A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: 18221, A NOVEL DUAL SPECIFICITY
; TITLE OF INVENTION: PHOSPHATASE AND
; FILE REFERENCE: 10448-030001
; CURRENT APPLICATION NUMBER: US/09/815,419
; CURRENT FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: US 60/191,858
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 654
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-815-419-3

Alignment Scores:
Pred. No.: 4,01e-19 Length: 654
Score: 256.00 Matches: 72
Percent Similarity: 41.00% Conservative: 35
Best Local Similarity: 27.59% Mismatches: 77
Query Match: 10.51% Indels: 77
DB: 10 Gaps: 5

US-09-761-640-4 (1-471) x US-09-815-419-3 (1-654)
QY 205 GlnValLeuHisGlnAlaCysGlnAlaAlaLeuGlySerGlyLeuValProGlyGlySer 224
   :|:|:|
DB 51 GAAGCAATGCACACAG-----GGTGAC 71
QY 225 AlaLeuThrThrPalaSerHisTyrGlnGluArgLeuAsnSerGluGlnSerCysLeuAsn 244
   ||| :|:|:|:|
DB 72 AACGCTAACTGGAAAGAAATTAAGAAACATGGAAGATGCCAGAAATTCATGTGT--- 128
QY 245 GluTyrThrAlaMetAlaAspLeuGlnSerLeuArgProProSerAlaGluProGlyGly 264
   :|:|:|:|
DB 129 -----GGA 131
QY 265 SerSerGluGlnGlnMetGlnAlaAlaIleArgAlaGluLeuTyr-LysValLeuAs 284
   ||| |||||
```


;; CURRENT APPLICATION NUMBER: US/09/816,494
;; CURRENT FILING DATE: 2001-03-23
;; PRIOR APPLICATION NUMBER: US 60/191,858
;; PRIOR FILING DATE: 2000-03-24
;; NUMBER OF SEQ ID NOS: 10
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO: 3
;; LENGTH: 1998
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-816-494-3

Alignment Scores:

Pred. No.:	2,67e-15	Length:	1998
Score:	227.50	Matches:	54
Percent Similarity:	58.39%	Conservative:	26
Best Local Similarity:	39.42%	Mismatches:	54
Query Match:	9,34%	Indels:	3
DB:	10	Gaps:	2

US-09-761-640-4 (1-471) x US-09-816-494-3 (1-1998)

QY 330 SerArgIlePheProHisLeuTyrLeuGlySerGluTrpAsnAlaAlaAsnLeuGlu 349
:::||||| |||:::|||||::: ||| |||
DB 478 ACCCGAATCTCCCATCTTATCTTGCGCCAGAGATGCTCCTCAAGAGAGCTG 537
QY 350 LeuGlnArgAsnArgValThrHisIleLeuAsnMetAlaArgGluIle-----AspAsn 367
:::|||||::: |||:::|||||::: |||
DB 538 ATGCAAGCAGATGGAGTGTATGTAAATGCCAGCATACCTGTCGAAGCCTGAC 597
QY 368 PheTyrProGluArgPheThrTyrHisAsnValArgLeuTrpAspGluIleSerAlaGln 387
||| ||| ||| ||| ||| ||| ||| ||| |||
DB 598 TTATCCCGAGCTCAT--TTCTGCGTGTGCTGTGAATGACAGCTTTGTGAGAAA 654
QY 388 LeuLeuProHisTrpLysGluThrHisArgPheIleGluAlaAlaArgAlaGlnGlyThr 407
:::|||||::: |||:::||||| |||::: |||
DB 655 ATTTGCCCTGCTGGACAATCAGTAGATTTCATTGAGAAAGCAAGCCTCCAAATGGA 714
QY 408 HisValLeuValHisCysLysMetGlyValSerArgSerAlaAlaThrValLeuAlaTyr 427
||||| ||| ||| ||| ||| ||| ||| ||| |||
DB 715 TGTGTTCTAGTGCACCTGTAGCTGATCCGCTCCGACCATCTGCTATCGCCTAC 774
QY 428 AlaMetLysGlnTyrGluGlySerLeuGlnGluAlaLeuArgHisValGlnGlnLeuArg 447
|||||::: ||| ||| ||| ||| ||| ||| ||| |||
DB 775 ATCATGAGAGAGATGACATGCTTTAGATGAAGCTTACAGATTGTGTGAAGAAAAAGA 834
QY 448 ProIleAlaArgProAsnProGlyPheLeuArgGlnLeuGlnIleTyrGln 464
||| ||| ||| ||| ||| ||| ||| ||| |||
DB 835 CCTACTATATCTCAAACTTCATTTTCTGGGCCAACTCTGACTATGAG 885

Search completed: April 12, 2003, 13:57:40
Job time : 371 secs